REMARKS

The invention relates to a transgenic mouse and a sperm cell comprising a mammalian retrotransposon.

Claims 34 through 49 are pending in the present application. Claims 1-33 have been canceled in a Preliminary Amendment dated September 1, 2000. Claims 35, 45 and 48 have been canceled in an Amendment filed on April 16, 2002. Therefore, claims 34, 36-44, 46, 47 and 49 currently pending and under examination.

Oath/Declaration

The Examiner has indicated that the Oath/Declaration filed March 23, 2005 is defective. Specifically, the Examiner states that the previous Declaration of John Moran does not make reference to a Preliminary Amendment that was filed September 1, 2000 in this divisional application.

Applicants submit herewith a newly executed declaration of John Moran.

Rejection of claims 34, 36-44, 46, 47 and 49 under 35 U.S.C. § 112, first paragraph

The Examiner has maintained the rejection of claims 34, 36-44, 46, 47 and 49 pursuant to 35 U.S.C. §112, first paragraph, for failing to comply with the enablement requirement. Specifically, the Examiner is of the opinion that the specification fails to provide an enabling disclosure for the claimed invention because the specification does not teach how to use the claimed transgenic mouse. The Examiner asserts further that the uses cited by the Applicants do not amount to a specific and substantial utility for the claimed transgenic mouse and sperm cell. Applicants respectfully submit that claims 34, 36-44, 46, 47 and 49 are enabled under 35 U.S.C. § 112 first paragraph for all of the reasons recited in the Responses to Office Actions dated April 16, 2002 and March 21, 2005, which are incorporated by reference herein in their entirety. Applicants make the following arguments.

The present invention is based on the discovery that an isolated DNAc molecule comprising a promoter and an L1 element can be introduced into a mouse in order to induce retrotransposition. Thus, the claimed invention encompasses both the transgenic mouse harboring the DNAc molecule and a sperm cell from the transgenic mouse. The Examiner is of the opinion that the specification fails to describe the transgenic mouse in such a way as to

enable one of ordinary skill in the art to make and use this invention. The use of the claimed transgenic mouse and sperm cell, set forth in the specification as filed and further elucidated below, reasonably correlate with the entire scope of the claim and therefore preclude the Examiner's rejection based on how to use the claimed invention.

At page 28, lines 14-16, the specification as filed provides the skilled artisan with enablement for using the transgenic mouse in elucidating animal and human gene function and evaluating targets for gene therapy. Specifically, an L1-expressing transgenic animal is useful in studying the biology of mammalian retrotransposition. As little is known about the process of retrotranposition in mammals and its effect on the mammalian genome, retrotransposition animal models, such as a transgenic mouse, are extremely useful in understanding this process and its impact on disease. Recent studies have in fact linked retrotransposition to disease by showing that spontaneous muscular dystrophy in a mouse model was caused by a retrotransposition event in the mouse laminin alpha2 chain gene (Besse et al. 2003, Neuromuscl. Disorders 13:216-222, a copy of which is enclosed herewith). Further, many L1 elements have reverse transcriptase activity (see specification, Example 3). As stated in the specification as filed on page 78, lines 7-14, reverse transcriptases have had a profound effect on the human genome, for example in the processing of Alu elements and pseudogenes for genome diversity. Therefore, in utilizing an L1-expressing transgenic animal to understand how retrotransposition occurs in mammals, it will provide necessary insight that is useful for studying human disease and may lead to potential therapeutics for treating these diseases.

Applicants also submit herewith a copy of Ostertag et al. (2002, Nature Genetics 32:655-660) which demonstrates that Applicants were able to generate a transgenic mouse which expresses an L1 element specifically in testis and ovary. Using this mouse model, Applicants were able to determine the frequency of the L1 insertions as well as the mutagenic potential of the L1 element in the animal. Applicants further submit herewith a copy of Prak et al. (2003, PNAS 100:1832-1837) which demonstrates that specific retrotransposition of the L1 element occurred in vivo, specifically in cultured sperm cells from a L1-EGFP (enhanced GFP) expressing transgenic mouse. These references also disclose non-lethal retrotransposon insertion sites in germ cells which have utility in the gene therapy field, for example, in determining potential vector insertion sites that are not harmful to the host subject. Taken together, these references provide support for Applicants assertion in the specification as filed that an L1

expressing animal model is useful in studying gene function and in evaluating targets for gene therapy.

Therefore, the enabled uses set forth by the Applicants reasonably correlate with the scope of the claimed invention. Further, the MPEP §2164.01 provides that if any uses are enabled when multiple uses are disclosed, the application is enabling for the claims. For the reasons discussed above, Applicants respectfully request that the rejection of the claims under 35 U.S.C. § 112, first paragraph, for lack of enablement, be reconsidered and withdrawn.

Summary

Applicants respectfully submit that each rejection of the Examiner to the claims of the present application has been overcome and that claims 34, 36-44, 46, 47 and 49 are now in condition for allowance. Reconsideration and allowance of these claims is respectfully requested at the earliest possible date.

Respectfully submitted,

Haig Kazazian *et al.*

June 30, 2006
(Date)

By: _____

Justin D.G. Brennan
Registration No. 52,650
DRINKER, BIDDLE & REATH, LLP
One Logan Square
18th and Cherry Streets
Philadelphia, PA 19103-6996
Telephone: (215) 988-2700
Direct Dial: (215) 988-2682
Facsimile: (215) 988-2757
E-Mail: Justin.Brennan@dbr.com
Agent for Applicants

JDGB

Enclosures: Communication
Ostertag et al. 2002, Nature Genetics 32:655-660
Prak et al. 2003, PNAS 100:1832-1837.
Besse et al. 2003, Neuormuscl. Disorders 13:216-222.
Request for Continued Examination
Declaration and Power of Attorney of John V. Moran
Five (5) Month Extension of Time

A mouse model of human L1 retrotransposition

Eric M. Ostertag[1], Ralph J. DeBerardinis[1], John L. Goodier[1], Yue Zhang[1], Nuo Yang[1], George L. Gerton[2] & Haig H. Kazazian Jr.[1]

Published online 4 November 2002; doi:10.1038/ng1022

The L1 retrotransposon has had an immense impact on the size and structure of the human genome through a variety of mechanisms, including insertional mutagenesis[1,2]. To study retrotransposition in a living organism, we created a mouse model of human L1 retrotransposition. Here we show that L1 elements can retrotranspose in male germ cells, and that expression of a human L1 element under the control of its endogenous promoter is restricted to testis and ovary. In the mouse line with the highest level of L1 expression, we found two *de novo* L1 insertions in 135 offspring. Both insertions were structurally indistinguishable from natural endogenous insertions. This suggests that an individual L1 element can have substantial mutagenic potential. In addition to providing a valuable *in vivo* model of retrotransposition in mammals, these mice are an important step in the development of a new random mutagenesis system.

L1 retrotransposons are directly or indirectly responsible for an estimated one-third of the human genome[1]. Of the approximately 500,000 L1 sequences in the human diploid genome, 40–80 remain active (that is, retrotranspositionally competent) and are a source of *de novo* mutations in humans[1–3]. Although a number of these recent L1 insertions have been characterized, it has not yet been possible to determine in which cells or when these events occur.

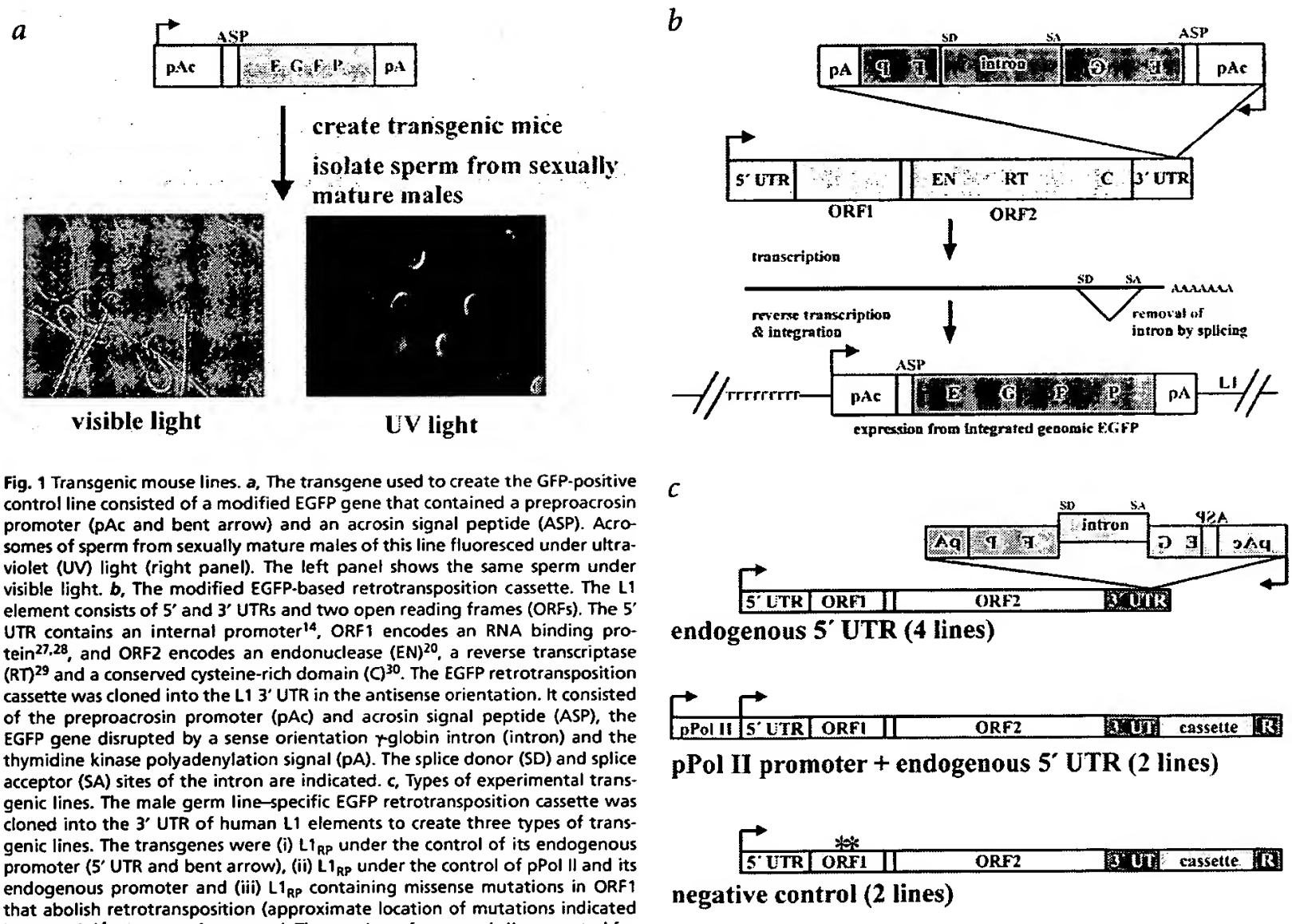
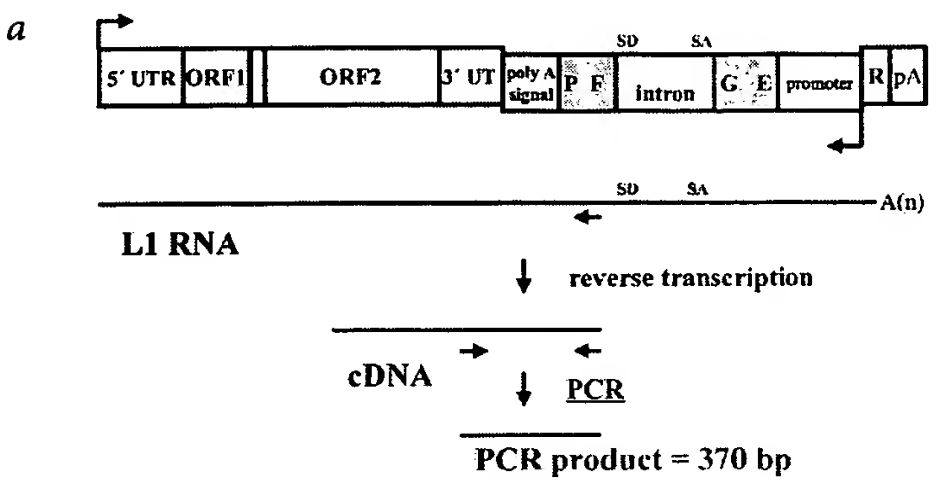


Fig. 1 Transgenic mouse lines. *a*, The transgene used to create the GFP-positive control line consisted of a modified EGFP gene that contained a preproacrosin promoter (pAc and bent arrow) and an acrosin signal peptide (ASP). Acrosomes of sperm from sexually mature males of this line fluoresced under ultraviolet (UV) light (right panel). The left panel shows the same sperm under visible light. *b*, The modified EGFP-based retrotransposition cassette. The L1 element consists of 5′ and 3′ UTRs and two open reading frames (ORFs). The 5′ UTR contains an internal promoter[14], ORF1 encodes an RNA binding protein[27,28], and ORF2 encodes an endonuclease (EN)[20], a reverse transcriptase (RT)[29] and a conserved cysteine-rich domain (C)[30]. The EGFP retrotransposition cassette was cloned into the L1 3′ UTR in the antisense orientation. It consisted of the preproacrosin promoter (pAc) and acrosin signal peptide (ASP), the EGFP gene disrupted by a sense orientation γ-globin intron (intron) and the thymidine kinase polyadenylation signal (pA). The splice donor (SD) and splice acceptor (SA) sites of the intron are indicated. *c*, Types of experimental transgenic lines. The male germ line–specific EGFP retrotransposition cassette was cloned into the 3′ UTR of human L1 elements to create three types of transgenic lines. The transgenes were (i) $L1_{RP}$ under the control of its endogenous promoter (5′ UTR and bent arrow), (ii) $L1_{RP}$ under the control of pPol II and its endogenous promoter and (iii) $L1_{RP}$ containing missense mutations in ORF1 that abolish retrotransposition (approximate location of mutations indicated by asterisks)[4], as a negative control. The number of transgenic lines created for each transgene type is indicated in parentheses.

[1]*Department of Genetics and* [2]*Center for Research on Reproduction and Women's Health, University of Pennsylvania School of Medicine, 475 Clinical Research Bldg., 415 Curie Blvd., Philadelphia, Pennsylvania 19104, USA. Correspondence should be addressed to H.H.K. (e-mail: kazazian@mail.med.upenn.edu).*



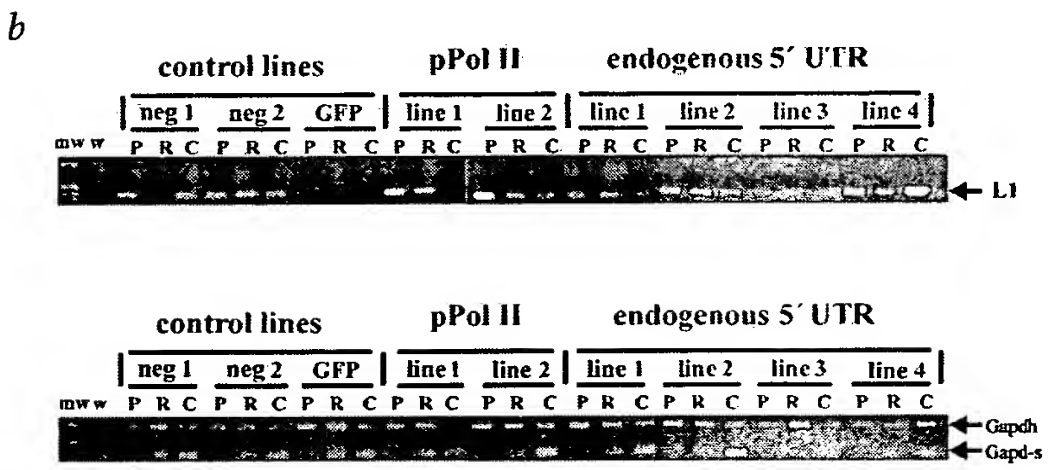


Fig. 2 L1 transgene expression in germ cell fractions. *a*, Strand-specific RT–PCR showing expression of the tagged L1 elements in spermatogenic fractions. We used an oligonucleotide primer specific for the EGFP gene to selectively reverse-transcribe transcripts originating from the L1 promoter. We then added a second oligonucleotide primer specific for the 3′ UTR of the L1 element to carry out PCR. A product of 370 bp is diagnostic of a tagged L1 transcript. *b*, RT–PCR results from the pachytene-spermatocyte (P), round-spermatid (R) and condensing-spermatid (C) fractions from all of our transgenic lines (upper panel). We did RT–PCR using mouse Gapdh primers as a control (lower panel). Lanes containing 1-kb molecular weight marker (mw) and RT–PCR of a water negative control (w) are indicated. P, pachytene spermatocytes; R, round spermatids; C, condensing spermatids.

A cultured-cell assay has been valuable for the study of L1 retrotransposition in a variety of transformed cell lines[4]. But cell culture experiments require retrotransposition from an episome to a chromosome, whereas natural retrotransposition occurs between chromosomal locations. In addition, natural retrotransposition in an organism probably differs from that in transformed cells in the factors used to control the timing and cell-type specificity of transcription as well as in other aspects.

Previous work on mouse L1 elements suggests that L1 expression is germ line–specific. Full-length, sense-strand L1 transcripts have been detected in prepuberal spermatocytes concomitantly with ORF1 protein[5]. Additionally, ORF1 was detected in germ cells of both male and female mice, Leydig cells of embryonic testis and theca cells of adult ovary[5,6] but not in normal somatic tissue. L1 retrotransposition has not, however, been demonstrated experimentally in any cell type in a living organism. Therefore, we created a mouse model of L1 retrotransposition to increase our understanding of L1 biology and to obtain an estimate of the impact of human L1 elements as insertional mutagens.

As a positive control, we created transgenic mice that express enhanced green fluorescent protein (EGFP) in the male mouse germ line and produce sperm with fluorescent acrosomes[7] (Fig. 1*a*). We then cloned an antisense intron into EGFP to create a retrotransposition cassette similar to that used in our cell culture assay[8] (Fig. 1*b*). EGFP transcripts arising from the marker's promoter before retrotransposition contain an antisense intron that cannot be spliced and, therefore, do not produce functional EGFP. A transcript from the L1 promoter that contains the antisense EGFP marker can be spliced, however, thereby removing the intron from the EGFP coding region. Cells express EGFP only after reverse transcription and integration of the intronless cDNA into chromosomal DNA. Therefore, the marker produces functional EGFP only after a retrotransposition event has occurred in male germ cells. We created three types of transgenic lines by cloning the cassette into the 3′ untranslated region (UTR) of three variants of $L1_{RP}$, the most active human element in cell culture[9,10](Fig. 1*c*): (i) $L1_{RP}$, (ii) $L1_{RP}$ with an additional promoter, the mouse RNA polymerase II large subunit promoter (pPol II) and (iii) JM111, an $L1_{RP}$ with two missense mutations in ORF1 that, unlike other L1 mutations characterized to date, completely abolish retrotransposition in cultured cells[4,8].

We purified pachytene spermatocytes, round spermatids and condensing spermatids from each of the mouse lines[11] and carried out strand-specific RT–PCR to verify expression of the L1 transgene (Fig. 2*a*). We detected expression of the tagged L1 transgene in all of the transgenic lines (Fig. 2*b*). We did RT–PCR using primers for mouse glyceraldehyde-3-phosphate dehydrogenase (Gapdh) to ensure that similar amounts of DNA were used in each reaction. Because the Gapdh promoter is silenced as spermatogenesis progresses, the Gapdh band intensity cannot be used to compare RNA amounts from one fraction to

Table 1 • Tissue distribution of L1 transgene expression

Line	Testis	Ovary	Kidney	Lung	Intestine	Liver	Brain
Negative control							
1	***	***	0	0	0	0	0
2	***	**	0	0	0	0	0
Endogenous 5′ UTR							
1	***	***	0	0	0	0	0
2	***	***	0	nd	0	0	0
3	***	**	0	0	0	0	0
4	***	***	0	0	0	0	0
pPol II							
1	***	***	**	**	**	*	*
2	***	nd	0	**	0	0	0

RT–PCR was done on RNA of various tissues of all transgenic lines. 0, no signal; asterisks, signal of varying intensity; nd, results of multiple experiments were equivocal.

another (for example, a pachytene-spermatocyte fraction to a round-spermatid fraction), but it can be used to compare within a fraction (for example, the pachytene-spermatocyte fraction from one line to the pachytene-spermatocyte fraction from another line). Notably, the Gapdh primers also amplify products from Gapd-s, a sperm-specific isoform of Gapdh that increases in abundance as spermatogenesis progresses[12]. Therefore, as Gapdh decreases in intensity, Gapd-s increases, providing an additional confirmation of the germ-cell fraction enrichment. In all RT–PCR experiments, we included RT-minus controls for which we did not detect bands, indicating that the RNA was not contaminated by DNA (data not shown).

In both of the pPol II lines, we found high expression of L1 in early spermatocyte fractions, which diminished in later fractions. This pattern corresponds well to the expression pattern of most mammalian promoters, which are silenced as spermatogenesis progresses[13]. In the lines driven by the L1 promoter alone, however, including the JM111 negative control lines, we found consistent expression throughout spermatogenesis. Notably, the human L1 promoter is expressed highly in male germ cells despite its differences from the various mouse L1 promoters[14–17], suggesting that similar *trans*-acting factors act on both promoters. Expression levels varied between transgenic lines and did not correlate with the estimated transgene copy number (data not shown). Strand-specific RT–PCR showed that transcription from the L1 promoter was restricted to the testis and ovary, whereas the pPol II promoter permitted transcription in other tissues (Table 1).

We first tested for the spliced EGFP mRNA that we expected to occur with retrotransposition by carrying out strand-specific RT–PCR on the condensing spermatid fractions (Fig. 3*a*). As expected, JM111 negative control lines showed no evidence of retrotransposition although they demonstrated expression from the transgene (Fig. 3*b*). In both lines driven by the pPol II promoter and one line (line 4) driven by the endogenous promoter, we detected retrotransposition. Those three lines expressed the highest levels of L1 RNA in pachytene spermatocytes, suggesting that the frequency of retrotransposition correlated with the strength of L1 expression in early germ-cell development. In this assay, functional EGFP transcript could only result from a double-stranded DNA containing the intronless EGFP marker, which could only exist if an EGFP-tagged L1 was transcribed, spliced and reverse-transcribed (that is, if a retrotransposition event had occurred). During L1 retrotransposition, the reverse transcription and chromosomal integration steps occur simultaneously during a coupled reaction termed target-primed reverse transcription[18]. We purified and sequenced several RT–PCR products. The products were of the expected sequence, confirming precise splicing of the intron.

To provide a rough estimate of the frequency of retrotransposition, we diluted RNA from the GFP control line into RNA from a JM111 negative control line to produce a dilution series. We used semi-quantitative RT–PCR to compare the experimental lines to the dilution series, and estimated that retrotransposition occurred in pPol II line 1 at around 1/1,000 spermatids and in other lines at somewhat lower frequencies. These are probably underestimates of retrotransposition frequencies because the amount of GFP RNA produced in a spermatid from a single-copy GFP retrotransposition event is probably lower than that produced from a multi-copy GFP transgene. This predicted reduction in RNA level is supported by the lower fluorescence of positive sperm from the pPol II line compared with positive sperm from the GFP-positive control line (Fig. 3*c*).

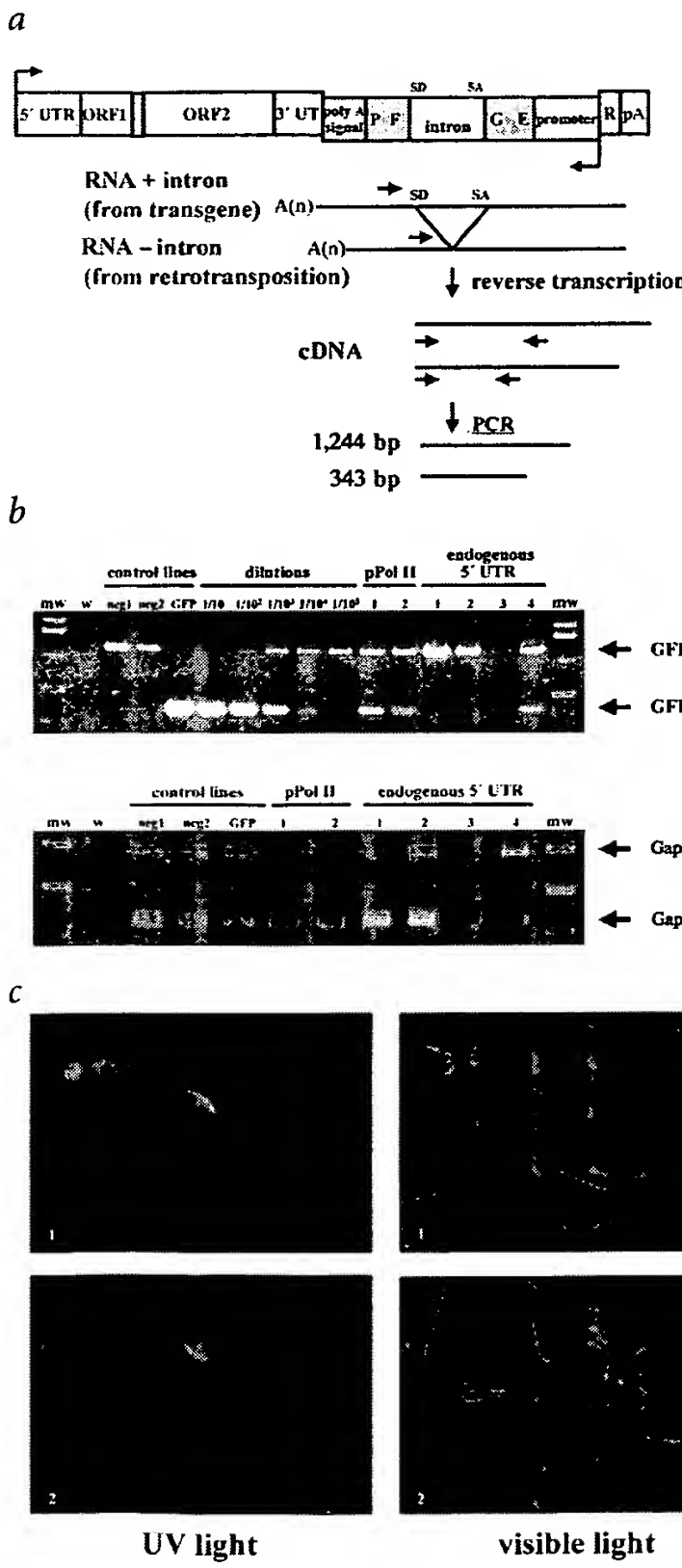


Fig. 3 Expression from retrotransposition events in condensing-spermatid fractions. *a*, We carried out strand-specific RT–PCR to detect EGFP expression from retrotransposition events. An oligonucleotide primer specific for EGFP was used to selectively reverse-transcribe RNA arising from the preproacrosin promoter. We expected two types of transcripts to arise from this promoter: those with an intron, which come from the transgene (1,244 bp), and those without an intron, which can only arise from a retrotransposition event (343 bp). We then did PCR after addition of an oligonucleotide primer flanking the intron. *b*, RT–PCR results from the condensing-spermatid fractions from all transgenic lines and the GFP-positive control line (upper panel). To obtain a minimum estimate of the retrotransposition frequency, we diluted condensing spermatid RNA from the GFP control line into condensing spermatid RNA from a negative control line to produce a dilution series. RT–PCR using mouse Gapdh primers was done as a control to ensure that similar amounts of DNA were used in each reaction (lower panel). Lanes containing a 1-kb molecular weight marker (mw) and RT–PCR of a water negative control (w) are indicated. *c*, Detection of retrotransposition by fluorescence microscopy. We isolated sperm from pPol II line1 and found that some sperm had fluorescent acrosomes when viewed under ultraviolet (UV) light (left panels). On the right are the same sperm under visible light. These pictures were taken with black and white film because it is more sensitive to low light than color film.

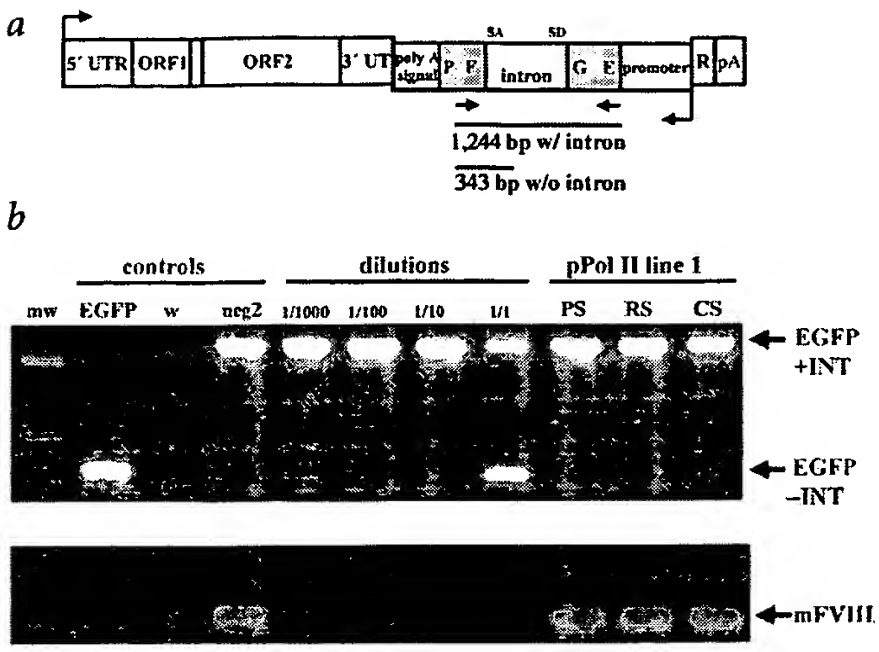


Fig. 4 PCR of germ-cell DNA to detect retrotransposition. *a*, The transgene included EGFP containing a γ-globin intron of 901 bp and would produce a band of 1,244 bp. A retrotransposition event would remove the intron and produce a band of 343 bp. *b*, We did PCR to detect EGFP on genomic DNA isolated from pachytene-spermatocyte (PS), round-spermatid (RS) and condensing-spermatid (CS) fractions from pPol II line 1 (upper panel). We created a dilution series by mixing plasmid DNA containing EGFP with genomic DNA from negative control line 2 (estimated copies of GFP per diploid genome are indicated). We did control PCRs on EGFP-containing plasmid (EGFP), water (w) and genomic DNA from negative control line 2 (neg2). A lane containing a 1-kb molecular weight marker is indicated (mw). We carried out a control PCR on neg2, PS, RS and CS genomic DNAs using mouse factor VIII primers to confirm that similar amounts of DNA were used in each reaction (lower panel).

To improve the estimate of the frequency of retrotransposition, we carried out PCR of genomic DNA from pPol II line 1 spermatogenic fractions (Fig. 4*a*). Using a dilution series consisting of plasmid DNA containing GFP sequences mixed with genomic DNA from a negative control line, we estimated that 1/100 spermatids or more contained a retrotransposition event (Fig. 4*b*).

Realizing that the RT–PCR and PCR results were only order-of-magnitude estimates of the frequency of retrotransposition, we sought formal proof that L1 retrotransposes at a high frequency in this mouse model by characterizing L1 insertions in the offspring of transgenic males of pPol II line 1. We bred F_2 and F_3 pPol II line 1 males and recovered 135 offspring. In these offspring, we found two retrotransposition events. Using inverse PCR, we characterized the insertions and found that they were typical of endogenous L1 retrotransposition events.

One insertion was 1.9 kb in length, contained a poly(A)$^+$ tail of 63 bp, had an inversion and was flanked by target-site duplications of 14 bp (Fig. 5*a*). The inversion occurred at the 5′ end and contained a deletion of 73 bp at the point of inversion. The last three nucleotides at the end of the 5′ target-site duplication were complementary to the nucleotides just proximal to the inversion point on the L1 RNA, suggesting that the inversion was produced by twin priming, a proposed mechanism for the creation of L1 inversions[19]. The predicted cleavage site was typical of L1 endonuclease sites, 5′-TTTT/AA-3′ (refs 20–22). This insertion segregated from the transgene in the subsequent generation.

The second insertion was 4.3 kb in length, contained a poly(A)$^+$ tail of 92 bp and was flanked by target-site duplications of 6 bp (Fig. 5*b*). Again, the predicted cleavage site was typical of that used by the L1 endonuclease. But the mouse that inherited this insertion did not inherit the transgene, not only indicating that retrotransposition almost certainly occurred from one chromosome to another but also strongly suggesting that the event occurred before the end of meiosis I. This result demonstrates that L1 elements can retrotranspose during male gametogenesis.

Our model shows high-frequency chromosome-to-chromosome retrotransposition of a human L1 element in the male germ line of an experimental animal. The *in vivo* insertions were indistinguishable structurally from endogenous L1 insertions in mammals. The observed retrotransposition frequency varied with the genomic context of the transgene and was highly correlated with the amount of L1 transcript present in the pachytene spermatocytes. In three of the four transgenic lines in which the endogenous promoter alone was used to drive L1 transcription, the level of expression in the pachytene spermatocytes was low, and the frequency of retrotransposition was below our limit of detection. In the fourth line, the level of expression in the pachytene spermatocytes was similar to that in the lines driven by the pPol II promoter, and the frequency of retrotransposition was roughly one-half that of pPol II line 1, as estimated by RT–PCR (Fig. 3*b*). We also found a correlation between L1 expression level and retrotransposition frequency in experiments using cultured cells (data not shown).

Among the 40–80 potentially active L1 elements in the human genome, some are older and common to all humans, whereas others are younger, and their presence is polymorphic. The younger elements are more active in cell culture[2,10,23], though most are far less active than $L1_{RP}$ owing to mutations that attenuate their ability to retrotranspose. The few very active elements are highly polymorphic. Therefore, each individual has a different complement of such L1 elements, only some of which reside in genomic regions that permit expression. This raises the question of how such a small number of potentially active L1 elements can impose such a high mutational load on the genome. Prior crude estimates of the frequency of L1 retrotransposition

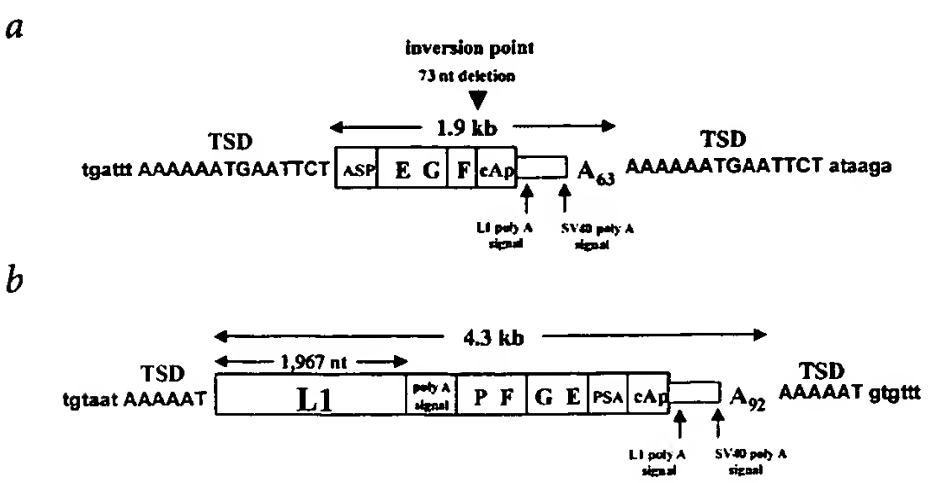


Fig. 5 L1 insertions in transgenic mice. *a*, Insertion #1 was 1.9 kb in length and contained an inversion of the 5′ end with a deletion of 73 bp at the inversion point, a common occurrence in L1 retrotransposition. A 63-bp poly(A)$^+$ tail was added after the SV40 poly(A)$^+$ signal, and the insertion was flanked by 14-bp target-site duplications (TSDs, sequence shown in capital letters). Flanking sequence is shown in lower-case letters. *b*, Insertion #2 was 4.3 kb in length and contained a 92-bp poly(A)$^+$ tail added after the SV40 poly(A)$^+$ signal, and the insertion was flanked by 6-bp target-site duplications (TSDs, sequence shown in capital letters). The L1 inserted into intron 1 of a predicted gene (mCG57584 from Celera Discovery System) on chromosome 9. Comparison of insertions #1 and #2 with their respective empty site sequences showed neither added nor deleted nucleotides other than the target-site duplications.

based on mutation rates in specific genes and overall mutation rates in germ cells ranged from 1 in 12 to 1 in 140 haploid genomes[24,25]. Our data suggest that a single very active human L1 can retrotranspose at a rate that is well within the range of previous empirical estimates when residing in a genomic locus that permits a high rate of transcription during spermatogenesis.

Our L1 retrotransposon mouse model will be valuable in answering a number of questions concerning the biology of mammalian retrotransposition. Is L1 retrotransposition in mice suppressed after multiple generations, analogous to retrotransposition of I factor in *Drosophila melanogaster*[26]? What are the factors that allow transcription and retrotransposition of L1 elements in the germ line, but suppress somatic cell transcription? Do defects in methylation or DNA double-strand–break repair increase the frequency of retrotransposition in germ cells? Human L1 transgenes may also be useful as random insertional mutagens for determining gene function in mice.

Methods

Cloning of transgenes. We created clone pBSKS-AcEGFP-INT (a male germ line–specific EGFP retrotransposition cassette) by carrying out a three-way ligation with an approximately 2,960-bp fragment from pBluescript KS(–) (Stratagene), a 453-bp *XhoI*/*SalI* fragment from pGES36-TKpolyA (a plasmid containing the EGFP retrotransposition cassette with a thymidine kinase poly(A)$^+$ signal) and a 2,395-bp *Hin*dIII/*Xho*I fragment from pRJD538. We cloned the AcEGFP cassette as an *Xma*I/blunted *Sal*I fragment into the 3′ UTR of $L1_{RP}$ (pJCC5-$L1_{RP}$ (ref. 8) cut with *Xma*I/*Bst*Z17I) to create pBS-$L1_{RP}$-AcEGFP or into the 3′ UTR of $L1_{RP}$(JM111), the $L1_{RP}$ element containing two missense mutations in the ORF1 coding region, (pJCC5-$L1_{RP}$(JM111)) cut with *Xma*I/*Bst*Z17I) to create pBS-$L1_{RP}$(JM111)-AcEGFP. We cloned the tagged $L1_{RP}$ from pBS-$L1_{RP}$-AcEGFP as a *Not*I/blunted *Apa*I fragment into the multiple cloning site of pRJD099 (promoter-less CEP4-based vector) cut with *Not*I/blunted *Sfi*I to create p99-$L1_{RP}$-AcEGFP or into the *Not*I/blunted *Apa*I–digested multiple cloning site of pRJD907 (CEP4-based vector containing the mouse pPol II promoter in place of the CMV promoter) to create p907-$L1_{RP}$-AcEGFP. We similarly cloned the tagged $L1_{RP}$(JM111) element from pBS-$L1_{RP}$(JM111)-AcEGFP into pRJD099 to create p99-$L1_{RP}$(JM111)-AcEGFP. We purified transgenes using Elutip-D mini-column (Schleicher & Schuell). The purified transgenes were microinjected into fertilized mouse oocytes (strain B6SJF1) by the University of Pennsylvania School of Medicine Transgenic & Chimeric Mouse Core Facility.

L1 expression assay. To detect expression of the tagged L1 transgene, we carried out a strand-specific RT–PCR with the OneStep kit (Qiagen) in 50 µl volume using 500 ng of RNA per reaction. We added oligonucleotide 1239+ during the reverse transcriptase step (30 min at 50 °C) to selectively reverse-transcribe transcripts arising from the L1 promoter (as opposed to the EGFP promoter) and then added oligonucleotide L16045(30) before the PCR step. Our PCR conditions included an initial step at 95 °C (15 min), 30 cycles of amplification (30 s at 94 °C, 45 s at 50 °C, 1 min at 72 °C) and a final step at 72 °C (10 min). We did RT-minus controls by setting up identical reactions in parallel and leaving the reactions on ice during the reverse transcription step. We used a PTC-200 Peltier Thermal Cycler (MJ Research) for RT–PCR. We digested samples for 30 min at 37 °C with 0.5 µl (250 ng) DNase-free RNase (Roche) and then analyzed samples by separating aliquots on a 1.0% agarose gel. We did control RT–PCRs using mouse Gapdh-specific oligonucleotides mGAPDH5′ and mGAPDH3′. We did 30 cycles of amplification (30 s at 94 °C, 45 s at 58 °C, 1 min at 72 °C) followed by a final step at 72 °C for 10 min. All primer sequences are available upon request.

Detection of retrotransposition events by RT–PCR. To detect expression from the EGFP gene of retrotransposition events, we carried out strand-specific RT–PCR as described above, except that we added oligonucleotide GFP1310R during the reverse transcriptase step to selectively reverse-transcribe transcripts arising from the EGFP promoter. We added oligonucleotide GFP968F before the PCR step. Our PCR conditions included an initial step at 94 °C (10 min), 35 cycles of amplification (10 s at 94 °C, 30 s at 66.8 °C, 30 s at 72 °C) and a final step at 72 °C (10 min). We did RT-minus controls by setting up identical reactions in parallel and leaving the reactions on ice during the reverse transcription step.

DNA sequence analysis. We used 35 ng of RT–PCR product and 3.2 pmol oligonucleotide primer GFP968F for sequencing reactions, which were done using ABI 377 and 373A Stretch sequencers (DNA sequencing core, University of Pennsylvania).

PCR of spermatogenic DNA. We isolated genomic DNA from spermatogenic germ-cell fractions with a Blood and Cell Culture DNA Mini Kit (Qiagen). To make a standard reference, we diluted pEGFP-N1 into round-spermatid fraction genomic DNA from negative control line 2 at 1/1,000, 1/100, 1/10 and 1 copies per haploid genome. We carried out PCR on 500 ng genomic DNA from all fractions with EGFP-specific primers GFP968F and GFP1310R. We did amplifications in 50 µl containing 1.25 U *Taq* DNA polymerase (Roche), 1× PCR reaction buffer (Roche), 0.2 mM of each dNTP, 200 ng of each oligonucleotide primer and approximately 500 ng genomic DNA. Our PCR conditions included an initial step at 94 °C (10 min), 35 cycles of amplification (10 s at 94 °C, 30 s at 66.8 °C, 30 s at 72 °C) and a final step at 72 °C (10 min). We did control PCR on genomic DNA from pachytene-spermatocyte, round-spermatid and condensing-spermatid fractions from pPol II mouse line 1 and control round-spermatid fraction genomic DNA from negative control mouse line 2 using mouse factor VIII-specific primers. The oligonucleotides used for PCR were MC-18 and MC-19. We did amplifications in 50 µl containing 1.25 U *Taq* DNA polymerase (Roche), 1× PCR reaction buffer (Roche), 0.2 mM of each dNTP, 200 ng of each oligonucleotide primer and approximately 500 ng genomic DNA. Our PCR conditions included an initial step at 94 °C (10 min), 35 cycles of amplification (10 s at 94 °C, 30 s at 58.4 °C, 1 min at 72 °C) and a final step at 72 °C (10 min). We analyzed samples by separating aliquots on a 1.0% agarose gel.

Transgenic mice. We bred all founder mice with non-transgenic 129/SV mice (Jackson Laboratories) to establish stable transgenic lines. In subsequent matings, we bred transgenic mice with non-transgenic 129/SV mice. The transgenic mouse protocol, titled "A human transposable element", was reviewed by the Institutional Biosafety Committee of the University of Pennsylvania and was approved on 28 April 2000. The protocol was reviewed by the Institutional Animal Care and Use Committee of the University of Pennsylvania and approved on 24 July 2000.

Characterization of insertions by inverse PCR. We digested genomic DNA with either *Sph*I (for insertion #1) or *Afl*II (for insertion #2) and carried out self-ligations on restricted fragments at low DNA concentrations. After ethanol precipitation, we subjected 400 ng of ligated DNA to an initial round of PCR using the primer set 1EGFPBRDB1 and 2EGFPDOWN with the Expand Long Template PCR System (Roche). We used a 2-µl aliquot from this reaction in a second-round nested PCR reaction using the primer sets 3EGFPUP and 5NEOGFPDOWNB for insertion #1 and 3EGFPUP and 4EGFPDWN1 for insertion #2. We isolated PCR products following gel electrophoresis. We obtained the sequence flanking the 3′ end of the L1 inserts using an oligonucleotide annealing at the end of the L1 poly(A)$^+$ tail, 5′ $A_{23}T$ 3′. We identified insertion sites for insertion #1 and #2 in entries GA_x5J8B7TT2F0 and GA_x5J8B7W3MYM, respectively, of the Celera Discovery System mouse genome database. Based on this sequence, we amplified insertion #2, together with about 320 nt of 5′ flanking DNA, using primers situated in the EGFP cassette (3EGFPUP) and flanking region (4AFL1AT5P2). We also amplified the entire insertion #1 using primers in the 5′ and 3′ flanking DNA, B141SPH3P2 and D141SPH5P4. We cloned PCR fragments into the vector pCR2.1 (Invitrogen) and sequenced them in their entirety.

Acknowledgments

We thank A. Farley, K.-S. Kim and M.C. Cha for technical assistance; the Transgenic and DNA Sequencing Cores of the University of Pennsylvania for generation of transgenic mouse lines and for DNA sequences, respectively; and J. Moran and B. Brouha for critical reading of the manuscript. E.M.O. was supported by a Howard Hughes Predoctoral Fellowship, and H.H.K. was supported by grants from the US National Institutes of Health.

Competing interests statement

The authors declare that they have no competing financial interests.

Received 18 April; accepted 24 September 2002.

1. The International Human Genome Sequencing Consortium. *Nature* 409, 860–921 (2001).
2. Ostertag, E.M. & Kazazian, H.H. Jr. Biology of mammalian L1 retrotransposons. *Ann. Rev. Genet.* 35, 501–538 (2001).
3. Sassaman, D.M. *et al.* Many human L1 elements are capable of retrotransposition. *Nature Genet.* 16, 37–43 (1997).
4. Moran, J.V. *et al.* High frequency retrotransposition in cultured mammalian cells. *Cell* 87, 917–927 (1996).
5. Branciforte, D. & Martin, S.L. Developmental and cell type specificity of LINE-1 expression in mouse testis: implications for transposition. *Mol. Cell. Biol.* 14, 2584–2592 (1994).
6. Trelogan, S.A. & Martin, S.L. Tightly regulated, developmentally specific expression of the first open reading frame from LINE-1 during mouse embryogenesis. *Proc. Natl Acad. Sci. USA* 92, 1520–1524 (1995).
7. Nakanishi, T. *et al.* Real-time observation of acrosomal dispersal from mouse sperm using GFP as a marker protein. *FEBS Lett.* 449, 277–283 (1999).
8. Ostertag, E.M. *et al.* Determination of L1 retrotransposition kinetics in cultured cells. *Nucleic Acids Res.* 28, 1418–1423 (2000).
9. Schwahn, U. *et al.* Positional cloning of the gene for X-linked retinitis pigmentosa 2. *Nature Genet.* 19, 327–332 (1998).
10. Kimberland, M.L. *et al.* Full-length human L1 insertions retain the capacity for high-frequency retrotransposition in cultured cells. *Hum. Mol. Genet.* 8, 1557–1560 (1999).
11. Romrell, L.J., Bellvé, A.R. & Fawcett, D.W. Separation of mouse spermatogenic cells by sedimentation velocity. A morphological characterization. *Dev. Biol.* 49, 119–131 (1976).
12. Welch, J.E., Schatte, E.C., O'Brien, D.A. & Eddy, E.M. Expression of a glyceraldehyde 3-phosphate dehydrogenase gene specific to mouse spermatogenic cells. *Biol. Reprod.* 46, 869–878 (1992).
13. Steger, K. Transcriptional and translational regulation of gene expression in haploid spermatids. *Anat. Embryol.* 199, 471–487 (1999).
14. Swergold, G.D. Identification, characterization, and cell specificity of a human LINE-1 promoter. *Mol. Cell. Biol.* 10, 6718–6729 (1990).
15. Loeb, D.D. *et al.* The sequence of a large L1Md element reveals a tandemly repeated 5' end and several features found in retrotransposons. *Mol. Cell. Biol.* 6, 168–182 (1986).
16. Padgett, R.W., Hutchison, C.A. 3rd & Edgell, M.H. The F-type 5' motif of mouse L1 elements: a major class of L1 termini similar to the A-type in organization but unrelated in sequence. *Nucleic Acids Res.* 16, 739–749 (1988).
17. Naas, T.P. *et al.* An actively retrotranposing, novel subfamily of mouse L1 elements. *EMBO J.* 17, 590–597 (1998).
18. Luan, D.D., Korman, M.H., Jakubczak, J.L. & Eickbush, T.H. Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposable elements. *Cell* 72, 595–605 (1993).
19. Ostertag, E.M. & Kazazian, H.H. Jr. Twin Priming, a proposed mechanism for the creation of inversions in L1 retrotransposition. *Genet. Res.* 11, 2059–2065 (2001).
20. Feng, Q., Moran, J.V., Kazazian, H.H. Jr. & Boeke, J.D. Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. *Cell* 87, 905–916 (1996).
21. Jurka, J. Sequence patterns indicate an enzymatic involvement in integration of mammalian retroposons. *Proc. Natl Acad. Sci. USA* 94, 1872–1877 (1997).
22. Cost, G.J. & Boeke, J.D. Targeting of human retrotransposon integration is directed by the specificity of the L1 endonuclease for regions of unusual DNA structure. *Biochem.* 37, 18081–18093 (1998).
23. Boissinot, S., Chevret, P. & Furano, A.V. L1 (LINE-1) retrotransposon evolution and amplification in recent human history. *Mol. Biol. Evol.* 17, 915–928 (2000).
24. Kazazian, H.H. Jr. An estimated frequency of endogenous insertional mutations in humans. *Nature Genet.* 22, 130 (1999).
25. Li, X. *et al.* Frequency of recent retrotransposition events in the human factor IX gene. *Hum. Mutat.* 17, 511–519 (2001).
26. Jensen, S., Gassama, M.P. & Heidmann, T. Taming of transposable elements by homology-dependent gene silencing. *Nature Genet.* 21, 209–212 (1999).
27. Hohjoh, H. & Singer, M.F. Cytoplasmic ribonucleoprotein complexes containing human LINE-1 protein and RNA. *EMBO J.* 15, 630–639 (1996).
28. Kolosha, V.O. & Martin, S.L. *In vitro* properties of the first ORF protein from mouse LINE-1 support its role in ribonucleoprotein particle formation during retrotransposition. *Proc. Natl Acad. Sci. USA* 94,10155–10160 (1997).
29. Mathias, S.L., Scott, A.F., Kazazian, H.H., Boeke, J.D. & Gabriel, A. Reverse transcriptase encoded by a human transposable element. *Science* 254, 1808–1810 (1991).
30. Fanning, T. & Singer, M. The line-1 DNA-sequences in 4 mammalian orders predicts proteins that conserve homologies to retrovirus proteins. *Nucleic Acids Res.* 15, 2251–2260 (1987).

Tracking an embryonic L1 retrotransposition event

Eline T. Luning Prak*†, Allen W. Dodson*, Evan A. Farkash*, and Haig H. Kazazian, Jr.‡

*Department of Pathology and Laboratory Medicine, University of Pennsylvania, 405B Stellar Chance Laboratories, 422 Curie Boulevard, Philadelphia, PA 19104; and ‡Department of Genetics, University of Pennsylvania, 475 Clinical Research Building, 415 Curie Boulevard, Philadelphia, PA 19104

Communicated by Martin G. Weigert, Princeton University, Princeton, NJ, December 13, 2002 (received for review June 15, 2002)

Long interspersed nuclear elements 1 (L1) are active retrotransposons that reside in many species, including humans and rodents. L1 elements produce an RNA intermediate that is reverse transcribed to DNA and inserted in a new genomic location. We have tagged an active human L1 element ($L1_{RP}$) with a gene encoding enhanced GFP (EGFP). Expression of GFP occurs only if L1-EGFP has undergone a cycle of transcription, reverse transcription, and integration into a transcriptionally permissive genomic region. We show here that L1-EGFP can undergo retrotransposition *in vivo* and produce fluorescence in mouse testis. The retrotransposition event characterized here has occurred at a very early stage in the development of an L1-EGFP transgenic founder mouse.

L1 transgenic mice | enhanced GFP

Long interspersed nuclear elements 1 (L1) are autonomous retrotransposons (1, 2). There are ≈40–80 active L1 elements in the human diploid genome (3, 4). Despite this rather unimpressive number, there are ≈500,000 inactive L1 elements, such that ≈17% of the mass of the genome is comprised of L1 sequences (5). The activity of L1s has contributed significantly to the shaping of the human genome. It is likely that L1 machinery has been "hijacked" by Alu elements, which are present in >1,000,000 copies per haploid genome (6, 7). L1s also contribute to processed pseudogene formation (8, 9). The demonstration that L1s can shuffle exons in cultured cells further highlights their evolutionary significance (10–12). Conversely, mammalian genomes may have co-opted the functions of transposable elements for their maintenance. For example, the enzyme telomerase has significant homology to the L1 reverse transcriptase (13).

Studying L1s is difficult because they are scattered in large numbers throughout the genome. To date, the activity of L1s *in vivo* has been inferred from the recovery of insertions that have been inherited and cause disease (14). Studies in the mouse have surveyed L1 RNA and ORF1 protein expression in various tissues (15, 16). However, until recently, a direct assay of L1 retrotransposition in mice has been lacking (17). We are interested in determining which tissues and cell types in the mouse are competent to support L1 retrotransposition. To approach this question, we took advantage of a cultured cell assay of retrotransposition (18). Instead of using neomycin phosphotransferase, we used enhanced GFP (EGFP) as a tag to track L1 insertions *in vivo* (17, 19). EGFP is expressed only when the L1 element undergoes a cycle of transcription, reverse transcription, and integration into a genomic site that permits EGFP transcription (Fig. 1*a*).

Here we document retrotransposition of the L1-EGFP transgene *in vivo*. On retrotransposition, the EGFP reporter is activated: green fluorescence is observed in the testes of mice that inherit the insertion. Genetic and phenotypic analysis of the mouse in which the insertion arose demonstrates that the insertion occurred at an early stage of development. These and other studies (17) indicate that tagged L1 transgenes may be used to study L1 biology *in vivo*. These L1-EGFP mice also illustrate an application for tagged L1 transgenes as cell lineage markers.

Materials and Methods

Transgene Construction. A fragment containing the entire $L1_{RP}$ retrotransposon (20) up to the unique *Bst*Z17I site at position 5964 in the 3′ UTR was cloned into a pBS (KS-) shuttle vector containing a genetically engineered L1.2 3′ UTR (18, 20). A linker (AAAAAAGTATACGTAAAAAAACCCGGGGGGA, gift from Eric Ostertag, University of Pennsylvania, Philadelphia) was cloned into the *Bst*Z17i and *Xma*I sites (which were preserved in the linker) with a unique *Sna*B1 site corresponding to position 5967 in the L1 3′ UTR (18). The EGFP marker [cytomegalovirus (CMV)-EGFPint-tkpA] was inserted via the *Sna*B1 and *Xma*I sites of the linker in JCC5-RPs-SnaB1. The RP-EGFP fragment was directionally cloned into a pCEP4 derivative (pol II-RJD99, a gift from Ralph DeBerardinis, University of Pennsylvania) that contained the mouse RNA polymerase II (large subunit) promoter (pol II) (21). A negative control construct was cloned by swapping part of RP with JM111. JM111 contains a human L1 element with mutations in ORF1 that abrogate retrotransposition activity (18). To clone JM111-RP-EGFP, JM111 and CMV-RP-EGFP (19) plasmids were isolated from dam-dcm- bacteria. This allowed us to use the (normally methylated) unique *Bcl*I site in L1 to swap the downstream portion of L1 and the EGFP marker from CMV-RP-EGFP into pJM111, creating a hybrid but nonfunctional L1 element. The transgene and negative control constructs were tested in a transient cell culture assay for EGFP expression (19). A similar construct has been used to make transgenic mice (JM111-EGFP) that lack genetically detectable retrotransposition events (17). The L1-EGFP transgene was liberated from the pCEP4 backbone by combined *Sal*I + *Mlu*I digestion and agarose gel electrophoresis, band purified (Gene Clean, Bio 101) followed by an Elutip column (Schleicher & Schuell). Three lines of L1-EGFP transgenic mice were produced (lines 57, 59, and 63). These transgenic mouse studies were approved by the University of Pennsylvania Institutional Animal Care and Use Committee.

PCR Primers. The primers, sequences, and uses are as follows: 6896rev, TATATCTCCCAATGCTATCC, inverse PCR (IPCR); SV40for, ATGATAAGATACATTGATGAGTTTGGA, IPCR; SV40rev, TCCAAACTCATCAATGTATCTTATCAT, RT-PCR; HSVtkrev, AAGGCGATGCGCTGCGAATCGG, RT-PCR; geno5, TTTATTGCCGATCCCCTCAGAAGAA, genotyping; geno3, TTCAAGGACGACGGCAACTACAAGA, genotyping; 6032for, AACACCCGTGCGTTTTATTC, intron PCR; and 6543rev, CAGCCCAGTTAGTCCTCTGC, intron PCR.

IPCR. Five to 10 μg of genomic DNA were digested for 4 h with *Ssp*I, extracted with phenol then chloroform, and subjected to overnight ligation in a 0.6-ml volume. The ligation products were reextracted, ethanol precipitated, and subjected to PCR amplification by using inverse IPCR primers situated just upstream of the SV40 polyA signal (SV40for) and just downstream of the EGFP cassette (6896rev). IPCR was carried out with the Expand Long kit using buffer system 1 following the manufacturer's directions (Roche, Nutley, NJ). The 4.3-kb amplicon was gel

Abbreviations: L1, long interspersed nucleotide element 1; EGFP, enhanced GFP; IPCR, inverse PCR; CMV, cytomegalovirus; MIE, major immediate early; SV, simian virus; pol II, polymerase II large subunit promoter.

†To whom correspondence should be addressed. E-mail: luning@mail.med.upenn.edu.

www.pnas.org/cgi/doi/10.1073/pnas.0337627100

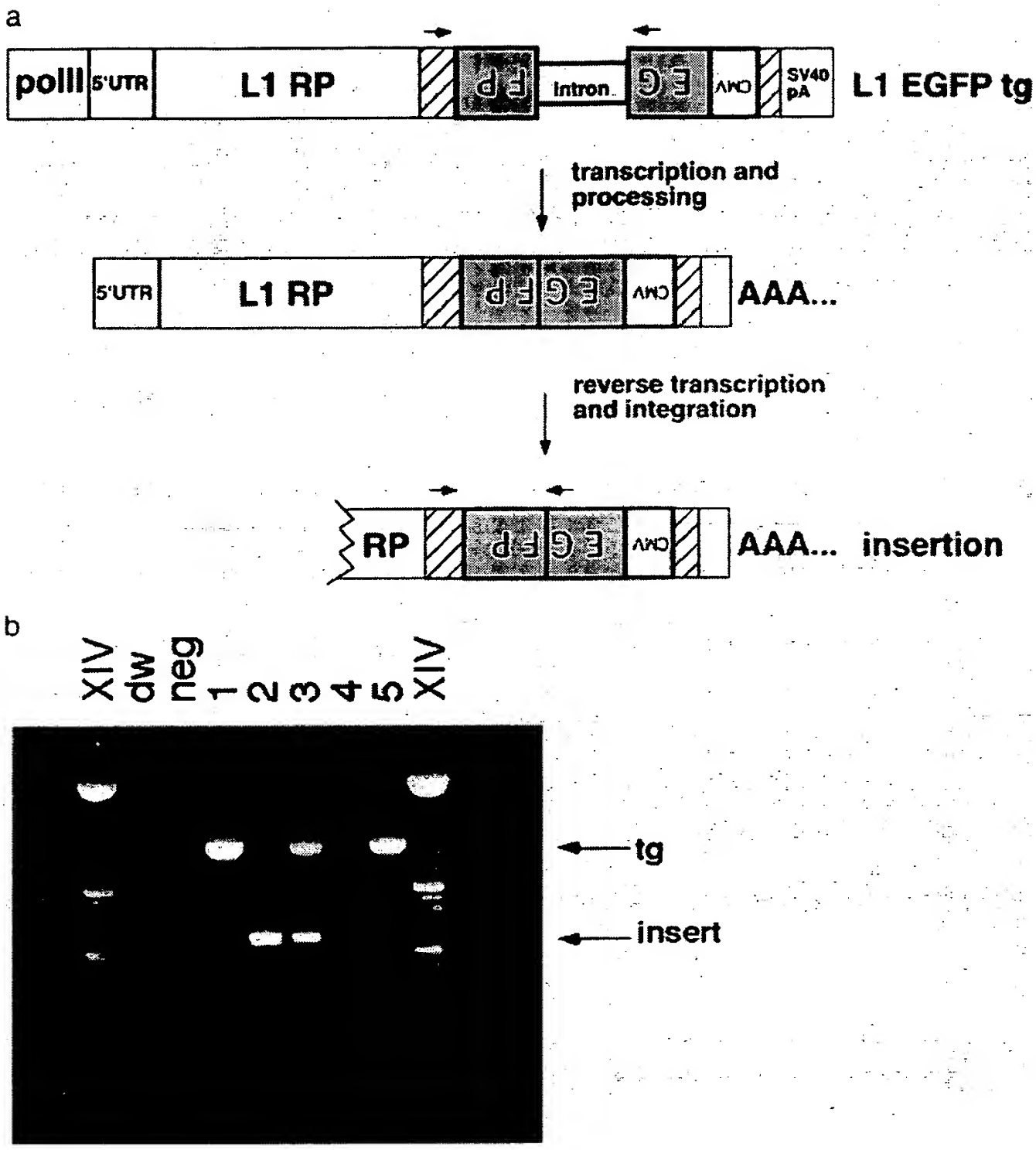


Fig. 1. Retrotransposition of L1-EGFP. (*a*) Schematic of the L1-EGFP transgene and its retrotransposition. L1 transcription is driven by the mouse RNA pol II promoter in addition to the L1 5′ UTR (21). The EGFP gene is in the antisense orientation relative to L1. EGFP (green) is situated in the 3′ UTR (hatched) of L1 and is interrupted by the mouse γ-globin intron. The intron is in the same transcriptional orientation as L1 (18). Therefore, when the L1 sense-strand transcript is processed, the γ-globin intron is spliced out. The EGFP gene is driven by the human CMV MIE promoter (pCMV-MIE) and has an HSV thymidine kinase polyadenylation sequence (tkpA). pCMV-MIE, EGFP, and tkpA are all antisense relative to $L1_{RP}$. At the very 3′ end of the L1-EGFP transgene is the SV40 late polyadenylation sequence (SV40pA) derived from the pCEP4 cloning vector (Invitrogen). 5′ truncation of the L1-EGFP insertion is depicted with a jagged line. Arrows depict the locations of the geno5 (left) and geno3 (right) genotyping primers used in the PCR assay shown in *b* (not drawn to scale). (*b*) The geno5 and geno3 primers flank the intron in EGFP and give rise to two products, a 1.5-kb amplicon (corresponding to the intron-containing transgene) and an ≈600-bp amplicon that lacks the 909-bp intron (corresponding to the insertion). Shown are the genotyping results on tail DNA from five offspring of founder 57 (lanes 1–5). dw, distilled water; neg, genomic DNA from the tail of a transgene negative mouse; tg, L1-EGFP transgene; XIV, 100-bp ladder with bright bands at 500 bp, 1,000 bp, and 2.6 kb (top band; Roche).

purified (Bio101; Qbiogene, Montreal) and cloned into pCR2.1 (Invitrogen). Portions of the cloned amplicon were sequenced (Fig. 2).

RT-PCR. RNA was extracted from mouse tissues using an RNAeasy midi kit (Qiagen, Chatsworth, CA) and digested with RNase-free DNAse1 (Roche). cDNA synthesis was performed with AMV reverse transcriptase following the manufacturer's instructions (Roche). The SV40rev oligonucleotide was used to produce L1 sense-strand cDNA (from the transgene), whereas the HSVtkrev oligonucleotide was used to produce EGFP cDNA (from the insertion). Each 50-μl PCR contained 5 μl of 10× PCR buffer (Roche buffer 1 with 15 mM $MgCl_2$), 1 μl of 10 mM dNTPs, 200 ng of each oligonucleotide, 0.3 μl of AmpliTaq Gold, and 5 μl of reverse transcriptase extension product. Amplifications were performed on a Peltier thermal cycler (Hybaid, Asford Middlesex, U.K.) by using the following program: 95°C for 15 min, followed by 40 cycles of 95°C for 25 sec, 62°C for 30 sec, and 72°C for 90 sec. This was followed by a final extension at 72°C for 10 min and a 4°C hold. The following tissues were surveyed for L1 sense-strand transcripts in L1-EGFP line 57 transgenic mice: testis, pooled ovaries, spleen, liver, and brain. Testis, pooled ovary, liver, and lung were surveyed in L1-EGFP line 59 transgenics. Testis and lung were surveyed in line 63 L1-EGFP transgenics. EGFP transcripts were surveyed in testis, spleen, liver, and brain of mice with the insertion.

```
1                                                                   60
aatatttacattacatggtagtacttgtttataagccaaaaaaaataaagaaatgattcat

61                                                                 120
ttaaatgtttaattaggatcccatattaaaatatgctactagaaataaaccaattaaata

121                                                                180
aaaccattaaatccaaatttatcatcataatccccccccccccccaatgtggcatttagc

181                  [mouse B2                                     240
tttaagaagcagtgtaaagagtctggacagatagctcagtggttaagagcactgactgtt

241                                                                300
cttctgagttcccagcaacgacatggtggctcacaatcatctgtaatgagatctgatgcc

301                                mouseB2]                        360
ctcttctggtgtgtctgaagacagcaacagtgtattcatacataaaaaaataaaataaat'

361                                                 5'TSD          420
aaataaataaatctttaaaaaaaaacagtataaaagttagttgttAAAAATAAATTGTTC

421                                                                480
TTCAAAAAACCAAACACCGCATATTCTCACTCATAGGTGGGAATTGAACAATGAGATCAC

481                              508
ATGGACACAGGAAGGGGAATATCACACT***

1                                                                   60
AACCCAACTAGAATGCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCT

61                             SV40 pA                             120
TTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAAAAAAAAAAAAAAAAAA

121                                                                180
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

181                                                 3' TSD         240
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAATTGTTCTaaac

241                  [mouse L1                                     300
acatgaaactcaagaaagaatgaagactgaagtgtggacactatgcccctccttagattt

301                                                                360
gggaacaaaacacccatggaaggagttacagagacaaagtttggagctgagatgaaagga

361                                                                420
tggaccatgtagagactgccatatccagggatccaccccataatcagcatccaaacgctg

421                                                                480
acaccattgcatacactagcaagattttattgaaaggacccagatgtagctgtctcttgt

481                                                                540
gagactatgccggggcctagcaaacacagaagtggatgctcacagtcagctaatggatgg

541                                                                600
atcatagggctcccaatggaggagctagagaaagtagccaaggagctaaagggatctgca

601                                                                660
accctataggtggaacaacattatgaactaaccagtaccccggaactcttgactctagct

661                                                                720
gcatatatatcaaaagatggcctagtcggccatcactggaaagagaggcccattggactt

721                                              mouse L1]         780
gcaaactttatatgccccagtacaggggaacaccagggccaaaaaggggggagtgggtggg

781                                  818
caggggagtgggggtgggtggatatgggggttctaggg
```

Fig. 2. Nucleotide sequence of the L1-EGFP insertion flanks. The genomic sequences flanking the L1-EGFP insertion are lowercase. L1-EGFP sequences adjoining the flanks are given in uppercase. The 5′ flanking sequence is shown in the top portion of the figure, followed by the 3′ flanking sequence in the bottom. Each flank is numbered separately. The complete L1-EGFP sequence (between the 5′ and the 3′ flanks) is not shown. The discontinuity corresponding to the L1-EGFP sequence is denoted by three asterisks at the end of the 5′ flank. Target site duplications (TSD) are in bold. The thymine at position 422, just downstream of the 5′ TSD, is not present in the $L1_{RP}$ sequence at that position. A nucleotide BLAST search of the mouse genome database using sequences flanking the L1-EGFP insertion suggests that the insertion is on *Mus musculus* chromosome 4 (http://www.ncbi.nlm.nih.gov).

Genotyping PCR. Genotyping PCR was carried out with the geno5 and geno3 primers, 250 ng of tail DNA, and a mix like the one described for RT-PCR. Amplification conditions were 94°C for 15 min, followed by 35–40 cycles of 94°C for 30 sec, 58°C for 30 sec, 72°C for 90 sec. This was followed by a final extension at 72°C for 5 min and a 4°C hold. Tail DNA from Fo57 offspring with the insertion band by genotyping PCR were separately amplified by using intron primers (and the same amplification conditions) to confirm the presence (or absence) of the transgene.

Primary Cell Cultures. A 5 × 5-mm portion of the inner aspect of the earlobe was harvested, washed three times in PBS, cut into smaller pieces, and compressed with the ribbed plunger of a syringe. Macerated tissue fragments were transferred to a sterile 15-ml conical tube, pelleted by centrifugation, and digested in 0.25% trypsin at 37°C for 1.5 h. The trypsin was neutralized with FCS, and the cells were pelleted and resuspended in 10 ml of fibroblast medium (DMEM with 4.5 g of glucose per liter and 10% FCS supplemented with penicillin and streptomycin). Cells were cultured in T25 flasks at 37°C in a humidified atmosphere with 7.5% CO_2. Tumor cells were isolated for culture by using a similar method, except that there were no washing steps before trypsin digestion, and trypsin digestion was carried out for ≈30 min.

Production of Ear Fibroblast Subclones. Fibroblasts were harvested by brief trypsinization, neutralized with serum, pelleted, and resuspended in fibroblast medium. The cell suspension was serially diluted and cultured in 96-well flat-bottom plates (Nunc). Cultures were evaluated for cell growth by light microscopy. Two groups of clones were used for analysis. One group consisted of eight subclones plated at a dilution at which at least two-thirds of the wells did not contain viable cells. A second group consisted of 11 wells that contained only one visible focus of cells. Lysates that typed negative by L1-EGFP genotyping PCR had amplifiable DNA by PCR for glyceraldehyde phosphate dehydrogenase.

Lysates for PCR Analysis. Ear fibroblast cultures and blastocysts were subjected to hypotonic lysis for 10 min at 80°C and overnight digestion with proteinase K at 55°C. Digested lysates were boiled for 15 min to inactivate the proteinase K and were stored at –20°C until PCR was performed. Five microliters of lysate were used in 25–50 μl of PCR.

Fluorescence Microscopy. Stereofluorescence microscopy was performed with a Leica MZFLIII stereofluorescent dissecting microscope, using a 100-W mercury bulb and the following filters: exciter HQ470/40 and emitter 515 nm LP (Chroma Technology, Brattleboro, VT). Images were captured by using a MagnaFire charge-coupled device (CCD) camera (Optronics International, Goleta, CA) and formatted in PHOTOSHOP (Adobe Systems, Mountain View, CA). Blastocysts were imaged for fluorescence at ×200–400 with the MagnaFire CCD camera mounted on an inverted fluorescence microscope (DM-IL, Leica, Deerfield, IL). Long exposures revealed a consistent pattern of autofluorescence among all of the blastocysts at a given developmental stage. The following tissues were surveyed by stereofluorescence in all three lines of L1-EGFP transgenic mice: liver, lung, heart, kidney, spleen, testis, ovary, uterus, skeletal muscle, and brain. In addition, thymus, bone marrow, eye, skin, and stomach were evaluated in line 57 L1-EGFP transgenics. The following organs were evaluated by fluorescence from mice with the L1-EGFP insertion: liver, lung, heart, kidney, thymus, spleen, testis, ovary, uterus, brain, eye, and skin.

Results

We chose a human L1 element, $L1_{RP}$, to study retrotransposition in mice, because $L1_{RP}$ is highly active, undergoing retrotransposition in approximately 1 of 30 cultured cells (20). Transgenes consisting of $L1_{RP}$ driven by its endogenous promoter, the 5′ untranslated region (5′ UTR), produced detectable transcripts only in the testis (E.T.L.P., unpublished data; ref. 17). Therefore, the mouse RNA polymerase II large subunit promoter (pol II) was added upstream of $L1_{RP}$ to achieve a broader tissue distribution of L1 expression (17, 21). The pol II promoter was added upstream of the 5′ UTR because constructs in which the 5′ UTR had been deleted were less active in the cell culture based retrotransposition assay (data not shown). The CMV-major immediate early (MIE) promoter was chosen to drive expression of EGFP because it was effective in producing fluorescent

Table 1. L1-EGFP genotyping data

Breeding surveys				
Cross	tg+	I+	tg+, I+	Negative
Fo57 × WT	13	8	4	17
57tg+/− × WT (or tg+)	64	0	0	41
59tg+/− × WT (or tg+)	54	0	0	26
63tg+/− (or 63Fo) × WT	13	0	0	24
Blastocyst (B) and morulae (M) surveys				
Fluorescence	Yes	No		
57tg+/− × WT	0	198		
Single B or M PCR	tg+	I+	tg+, I+	Negative
57tg+/− × WT	17	0	0	18
Ear fibroblast subclones from Fo57				
Clones	tg+	I+	tg+, I+	Negative
19	9	0	3	7

Three independent lines of L1-EGFP mice (57, 59, 63) were bred to screen for germ-line insertion events by genotyping PCR of tail DNA. All transgenes were maintained on the male line beyond the first generation. Fo57, line 57 founder; 129/SvEmsJ, 57tg+/−, heterozygous for the line 57 L1-EGFP transgene; tg+, transgene positive; I+, has the L1-EGFP insertion.

signals in the cultured cell assay and was known to be expressed in a wide range of tissues (19, 22). Transgene constructs were tested for retrotransposition activity in cultured HeLa cells before pronuclear injection. No EGFP-expressing cells were observed with a negative control construct that contained a nonfunctional L1 element (JM111-EGFP). To control for integration site effects on transgene expression and other unknown position effects on retrotransposition, we generated three independent lines of L1-EGFP transgenic mice.

PCR genotyping of offspring from an L1-EGFP transgenic founder (founder number 57, Fo57) revealed a low molecular-weight amplicon that was consistent in size with a spliced L1-EGFP fusion and a high molecular-weight amplicon that was consistent with the original transgene (Fig. 1*b*). DNA sequence analysis confirmed the identities of both bands.

To verify that the spliced L1-EGFP sequence arose via retrotransposition, we cloned the genomic flanks of the L1 insertion by inverse PCR. The insertion is a truncated L1, consisting of 208 base pairs of the $L1_{RP}$ ORF2 and the 3′ UTR containing the entire EGFP cassette lacking the intron (Fig. 2). L1 insertions often truncate within the most downstream kilobase of the L1 sequence (23). This particular insertion is flanked by target site duplications having the sequence AAAAATA-AATTGTTCT. Genomic L1 target sites are characteristically 7–20 nt in length and AT rich (2). The presence of target site duplications is consistent with an endonuclease-dependent retrotransposition event. In the 3′ flank, the target site is preceded by a polyadenylate tail, which originates from the SV40pA signal. The 5′ genomic flank contains part of a mouse B2 element, whereas the 3′ genomic flank contains 506 bases of the 3′ end of a mouse L1 element (Fig. 2).

Of 42 offspring from Fo57, 13 inherited the transgene, 8 inherited the insertion, 4 inherited both, and 17 inherited neither (Table 1). Thus, the insertion segregates from the transgene in offspring of Fo57, indicating that it is likely to be at a significant physical distance from the transgene. Frequent transmission of the insertion could be due to multiple independent insertions in the germ cells of the founder or to a single insertion that took place in a cell that ultimately gave rise to ≈30% of germ cell precursors. To distinguish between these two alternatives, we performed Southern analysis of genomic DNA from tail biopsies of four different offspring harboring the insertion-sized PCR amplicon. When probed with EGFP exonic sequence, all of the mice with the insertion shared the same sized band (data not shown.) This result was confirmed with a second restriction enzyme and implies a single retrotransposition event.

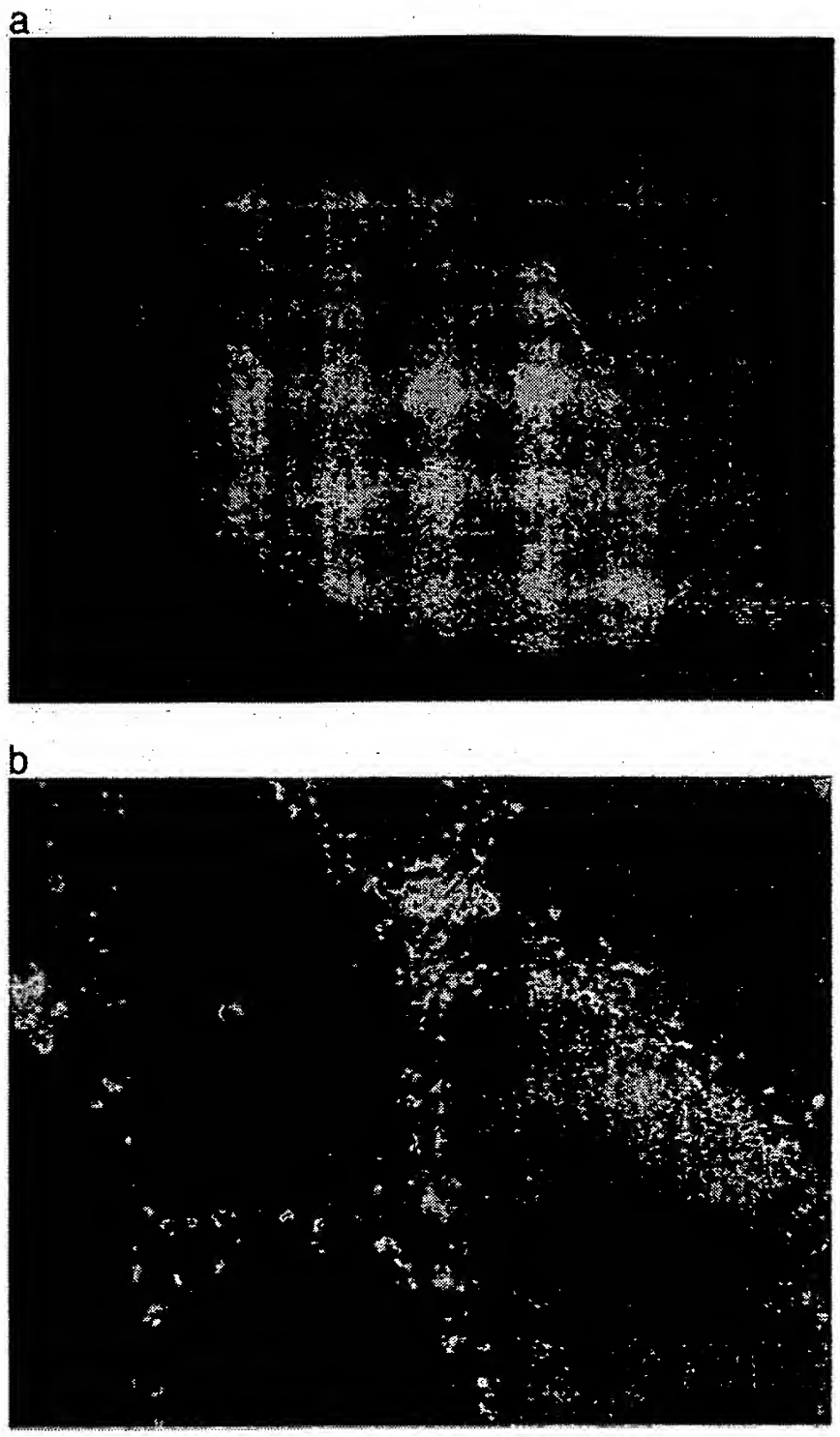


Fig. 3. EGFP fluorescence in the seminiferous tubules of Fo57. (*a*) Stereofluorescent image (×20), testis (fresh tissue) from Fo57. (*b*) Fluorescent light image (×200, frozen section from Fo57 testis. GFP fluorescence (dark green) is restricted to the seminiferous tubule in the lower right. Interstitial areas are autofluorescent (yellow–green speckled pattern).

At the age of 11.5 mo, Fo57 was killed because of a rapidly growing 9-g tumor originating from the right leg/hip. On the basis of histology, we believe that the tumor is a high-grade sarcoma, possibly a rhabdomyosarcoma (data not shown). Very few tumors and no sarcomas have been reported in the strain of Fo57, (SJL/B6)F1 (http://tumor.informatics.jax.org). The freshly harvested tumor was not fluorescent, nor was GFP fluorescence detected in frozen sections of the tumor. However, tumor DNA contained both the transgene and the insertion by PCR (data not shown). We have intercrossed mice carrying the line 57 transgene and have not seen any tumors in offspring (now up to 6 mo of age), including those that are homozygous for the transgene. We have also intercrossed mice carrying the insertion and have not observed any tumors to date in the homozygous offspring (now 6 mo of age). Given the late age of tumor onset in the founder and the absence of

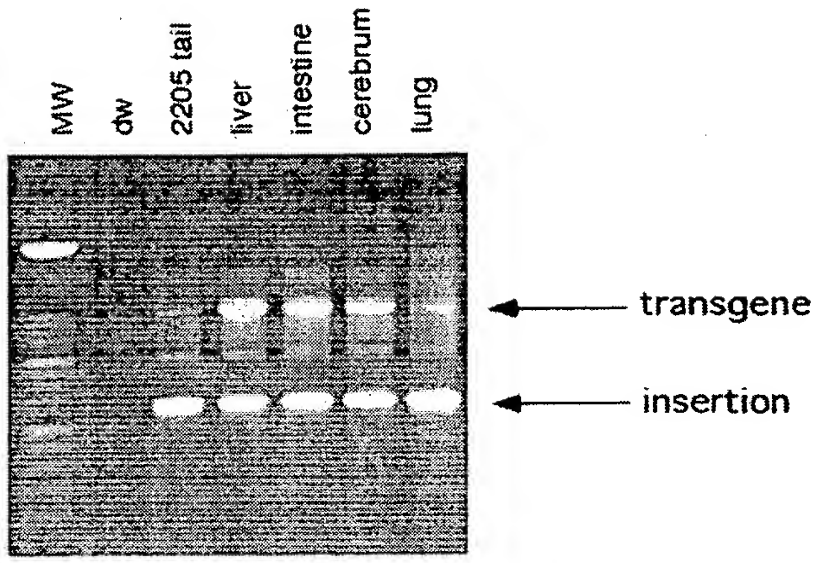


Fig. 4. Tissue DNA PCR surveys in Fo57. Genomic DNA from Fo57 tissues (liver, intestine, cerebrum, and lung) was amplified with geno5 and geno3 primers. A short extension time was used to favor the smaller amplicon (corresponding to the L1-EGFP insertion) relative to the larger amplicon (corresponding to the transgene). MW, 100-bp molecular weight ladder; dw, water; 2205, tail DNA from a line 57 mouse that has the transgene (present in one to five copies) and the single copy insertion.

similar tumors in other line 57 mice with the transgene or the insertion, the development of this tumor is likely unrelated to the transgene or the L1 insertion.

The only organ with demonstrable fluorescence in Fo57 was the testis (Fig. 3*a*). On frozen section, GFP was discernible in individual seminiferous tubules (Fig. 3*b*). This pattern of uniform fluorescence across longitudinal aspects of single tubules suggests that cells giving rise to segments within the tubules already contained the L1-EGFP insertion (24). Given the evidence for a single insertion event among offspring inheriting the insertion, this pattern of fluorescence suggests that these segment-producing cells were derived from an earlier precursor cell. To determine whether the precursor cell gave rise to any tissues other than the testis, Fo57 genomic DNA was extracted from the liver, small intestine, cerebrum, and lung and amplified for L1-EGFP. Amplicons corresponding to the transgene and the insertion were obtained in all of these tissues (Fig. 4). It is possible that only a few cells or even a single lineage of cells harboring the insertion could account for these results. Alternatively, metastatic spread of the tumor could have seeded these different organs. However, we disfavor the theory that the insertion arose in a tumor cell, because it is inconsistent with the pattern of fluorescence seen in the testis, which also shows no gross or microscopic evidence of tumor (Fig. 3).

To assess mosaicism in Fo57 directly, we cultured ear fibroblasts and produced monoclonal subclones by limiting dilution. Of 19 subclones, 7 were negative for both the transgene and the insertion (but contained amplifiable DNA), 9 had the transgene, and 3 had both the transgene and the insertion (Table 1). These results confirm that Fo57 is mosaic for the transgene and the insertion. They also indicate that the L1-EGFP retrotransposition event occurred at an early stage of development.

The L1-EGFP insertion resulted in heritable fluorescence that was restricted to the testis. Fig. 5 shows that the testis from a mouse that inherited the insertion from Fo57 is fluorescent, whereas testes from WT littermates are negative. Sperm were not fluorescent, possibly due to cytoplasmic loss during spermatogenesis or lack of EGFP expression in this cell type (25). We were surprised that fluorescence was limited to the testis in mice that had inherited the insertion, given the broad pattern of expression of the CMV-MIE promoter in transgenic mice (22). EGFP sense-strand RNA was readily detected by RT-PCR in the testis and was present at lower levels in

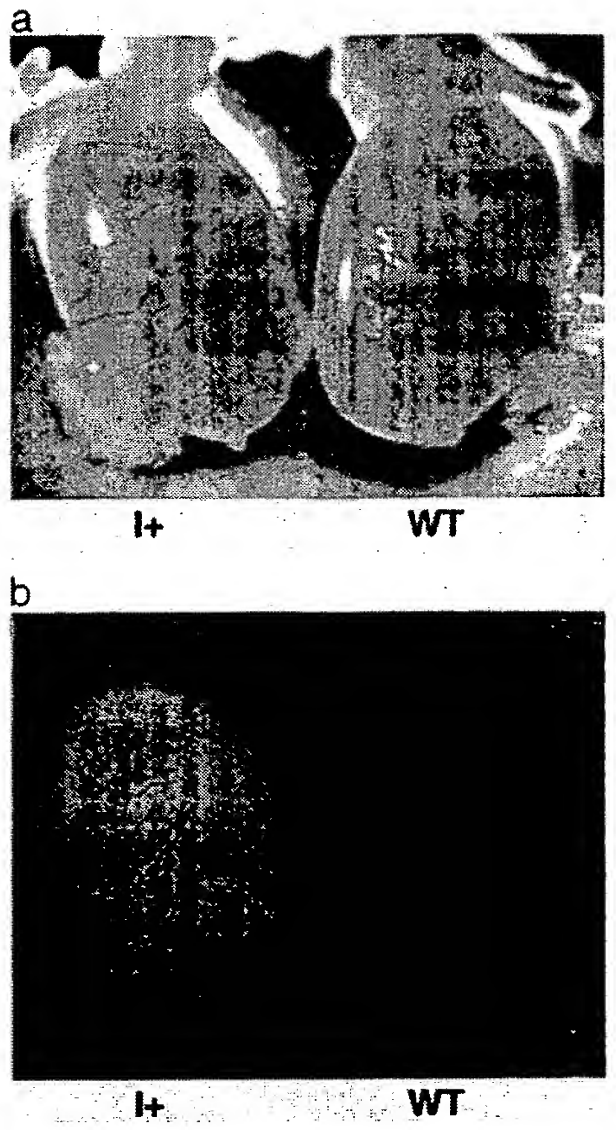


Fig. 5. EGFP fluorescence in the testes of mice inheriting the L1-EGFP insertion. (*a*) Transmitted light image, testis from an F1 mouse with only the insertion (I+, son of Fo57), and a WT littermate. (*b*) Stereofluorescent light image, same testis pair as in *a*.

multiple organs outside of the testis (data not shown). Lower levels of EGFP transcripts in organs outside of the testis suggest that decreased or undetectable fluorescence could be explained by decreased transcript abundance. DNA sequence analysis revealed no mutations in the CMV promoter in the insertion.

We searched for somatic retrotransposition events by surveying numerous organs from all three lines of L1-EGFP mice for EGFP fluorescence. All organs surveyed were negative (see *Materials and Methods*). These findings suggest that the frequency of somatic retrotransposition of the L1-EGFP transgene is low. One possible explanation for the low frequency is that the transgene or the marker gene is rarely expressed or not readily detected in somatic tissues. Consistent with this explanation, sense-strand L1 RNA was low or undetectable in organs other than the testis in several independent RT-PCR experiments (data not shown).

Given the low levels of L1 transcripts in organs other than the testis, we focused on looking for germ-line retrotransposition events in transgenic animals. Of 222 offspring from three independent lines of L1-EGFP transgenic mice, 131 were transgene positive, 91 were transgene negative, and none exhibited a spliced L1-EGFP product by PCR (Table 1). We also assayed 56 morulae and 142 blastocysts derived from mating a heterozygous transgene positive line 57 male to superovulated WT females. None of the blastocysts or morulae was fluorescent. It is possible that the EGFP marker is not expressed in blastocysts. Therefore, we looked for retrotransposition events by PCR of single blastocyst lysates with primers flanking the EGFP intron. Of 35

lysates surveyed, approximately half were transgene positive, and none contained the smaller spliced product (Table 1). Taken together, these findings lead us to a conservative estimate of an L1-EGFP germ-line retrotransposition frequency of <1 in 100.

Discussion

We have traced the origins of an L1 retrotransposition event to an early stage in the embryonic development of a mouse, Fo57. Fo57 is mosaic for the insertion as well as the L1-EGFP transgene. This is demonstrated visually for the L1 insertion by expression of EGFP in some but not all tubules within the Fo57 testis (Fig. 3*a*). Mosaicism is also supported by our limiting dilution analysis of skin fibroblasts of Fo57, in which we recover cells that are either transgene positive, transgene and insertion positive, or negative for both (Table 1). These three different genotypes indicate that the insertion took place after the two-cell stage, possibly as early as the four-cell stage.

Little is known about the activity of L1 elements during the early stages of mouse development. Others have shown that full-length L1 transcripts can be found in murine blastocysts (26). L1 ORF1 protein and sense-strand L1 RNA are expressed in male and female murine germ cells and embryos (15, 16). Recently L1s have been demonstrated to retrotranspose in male germ cells (17). Retrotransposition in germ cells is an attractive strategy for a genomic parasite. In theory, germ cells could tolerate a high rate of retrotransposition without significantly reducing the reproductive fitness of the host, but the chances of transmitting a particular insertion to the offspring are low. Here we provide evidence for another approach for L1 to colonize the genome: embryonic retrotransposition. An L1 retrotransposon that inserts in an embryonic cell (or germ cell precursor) has greatly increased its chances of propagating that insertion to the next generation, unless the insertion is deleterious to the host.

It is possible that the L1-EGFP insertion observed here originated from an extrachromosomal source. Clearly it is mechanistically allowable for retrotransposition to proceed from an extrachromosomal substrate, such as a plasmid in the cultured cell assay (18). Here we show that the cellular machinery exists for an L1 of unknown structure (extrachromosomal or chromosomal) to retrotranspose during early development.

We have been unable to identify any germ-line L1 retrotransposition events in any of the 222 progeny and 198 blastocysts derived from our L1-EGFP transgenic mice. These results differ significantly from the retrotransposition frequency of 1 in 70 sperm observed by Ostertag *et al.* (17) in one of their L1-EGFP transgenic lines. Chromosomally based L1s or multicopy transgene arrays may be subjected to a variety of constraints that could limit L1 mobility, including transcriptional inhibition by methylation or posttranscriptional silencing (27). It is possible, for example, that the integration sites of our transgenes were unfavorable compared with the most active line in Ostertag *et al.* Alternatively, the difference in rate could be due to differences in the transgenes themselves, such as the promoters driving the EGFP marker gene. EGFP is likely to be transcribed in both the sense and antisense directions (relative to L1) in our L1-EGFP transgenic mice. The EGFP transcript, originating from the CMV-MIE promoter, is expressed in multiple tissues and may be active at a very early stage of development, unlike the acrosin promoter used by Ostertag *et al.* (17), which appears to be restricted to the male germ line (22, 28). The EGFP transcript (which is antisense relative to L1) may inhibit L1 transcription or form dsRNA that targets other L1-EGFP transcripts for destruction.

L1-EGFP transgenics may prove useful in determining whether altered host cell characteristics (such as tissue injury, neoplastic transformation, defective DNA repair machinery, or altered RNA regulation) result in increased L1 mobility. If L1s can be induced to retrotranspose at high frequencies, they could also be used as cell lineage markers. The L1 insertion permanently tags the cell in which it arose and all of the descendants of that cell. As L1 insertions accumulate, a dendogram of lineal descendants could be created.

We thank Julia Sarino and Stephanie Johnson for technical assistance; Paula Stein for the blastocyst harvests; the University of Pennsylvania Transgenic facility; and the Institute for Human Gene Therapy morphology core facility. We thank George Gerton, Brad Johnson, Shane Horman, James Wheeler, Paul Zhang, Richard Schultz, and members of the Kazazian laboratory for helpful discussions and cloning reagents. E.T.L.P. is supported by the National Cancer Institute (5-K08-CA83977), the University of Pennsylvania Research Foundation, and the McCabe Foundation. A.W.D. was supported by a summer fellowship from Biomedical Graduate Studies at the University of Pennsylvania. E.A.F. is supported by the National Institutes of Health (NIH) (T32 H10791) and H.H.K. is also supported by the NIH.

1. Skowronski, J., Fanning, T. G. & Singer, M. F. (1988) *Mol. Cell. Biol.* **8,** 1385–1397.
2. Kazazian, H. H. & Moran, J. V. (1998) *Nat. Genet.* **19,** 19–23.
3. Ostertag, E. & Kazazian, H. H., Jr. (2001) *Annu. Rev. Genet.* **35,** 501–538.
4. Sassaman, D. M., Dombroski, B. A., Moran, J. V., Kimberland, M. L., Naas, T. P., DeBerardinis, R. J., Gabriel, A., Swergold, G. D. & Kazazian, H. H., Jr. (1997) *Nat. Genet.* **16,** 37–43.
5. Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., *et al.* (2001) *Nature* **409,** 860–921.
6. Boeke, J. D. (1997) *Nat. Genet.* **16,** 6–7.
7. Smit, A. F. A. (1999) *Opin. Genet. Dev.* **9,** 657–663.
8. Esnault, C., Maestre, J. & Heidmann, T. (2000) *Nat. Genet.* **24,** 363–367.
9. Wei, W., Gilbert, N., Ooi, S. L., Lawler, J. F., Ostertag, E. M., Kazazian, H. H., Boeke, J. D. & Moran, J. V. (2001) *Mol. Cell. Biol.* **21,** 1429–1439.
10. Moran, J. V., DeBerardinis, R. J. & Kazazian, H. H., Jr. (1999) *Science* **283,** 1530–1534.
11. Pickeral, O. K., Makalowski, W., Boguski, M. S. & Boeke, J. D. (2000) *Genome Res.* **10,** 411–415.
12. Goodier, J. L., Ostertag, E. M., Kazazian, H. H., Jr. (2000) *Hum. Mol. Genet.* **9,** 653–657.
13. Eickbush, T. H. (1997) *Science* **277,** 911–912.
14. Kazazian, H. H., Jr. (1998) *Curr. Opin. Genet. Dev.* **8,** 343–350.
15. Trelogan, S. A. & Martin, S. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 1520–1524.
16. Brancifort, D. & Martin, S. L. (1994) *Mol. Cell. Biol.* **14,** 2584–2592.
17. Ostertag, E., DeBerardinis, R. J., Goodier, J. L., Zhang, Y., Yang, N., Gerton, G. & Kazazian, H. H. (2002) *Nat. Genet.* **32,** 655–660.
18. Moran, J. V., Holmes, S. E., Naas, T., DeBerardinis, R. J., Boeke, J. D. & Kazazian, H. H. (1996) *Cell* **87,** 917–927.
19. Ostertag, E. M., Luning Prak, E. T., DeBerardinis, R. J., Moran, J. V. & Kazazian, H. H. (2000) *Nucleic Acids Res.* **28,** 1418–1423.
20. Kimberland, M. L., Dovoky, V., Prchal, J., Schwahn, U., Berger, W. & Kazazian, H. H. (1999) *Hum. Mol. Genet.* **8,** 1557–1560.
21. Ahearn, J. M., Bartolomei, M. S., West, M. L., Cisek, L. J. & Corden, J. L. (1987) *J. Biol. Chem.* **262,** 10695–10705.
22. Baskar, J. F. Smith, P. P., Ciment, G. S., Hoffman, S., Tucker, C., Tenney, D. J., Colberg-Poley, A. M., Nelson, J. A. & Ghazal, P. (1996) *J. Virol.* **70,** 3207–3214.
23. Boissinot, S., Chevret, P., Furano, A. V. (2000) *Mol. Biol. Evol.* **6,** 915–928.
24. *Histological and Histopathological Evaluation of the Testis* (1990), eds. Russell, L. D., Ettlin, R. A., Hikim, S. & Clegg, E. D. (Cache River Press, Clearwater, FL), pp. 41–58.
25. Villuendas, G. Gutierrez-Adan, A., Jimenez, A., Rojo, C., Roldan, E. R. & Pintado, B. (2001) *Int. J. Androl.* **24,** 300–305.
26. Packer, A. I., Manova, K. & Bachvarova, R. F. (1993) *Dev. Biol.* **157,** 281–283.
27. Bestor, T. H. & Tycko, B. (1996) *Nat. Genet.* **12,** 363–367.
28. Watanabe, K., Baba, T., Kashiwabara, S., Okamoto, A. & Arai, Y. (1991) *J. Biochem. (Tokyo)* **109,** 828–833.

PERGAMON

Neuromuscular Disorders 13 (2003) 216–222

www.elsevier.com/locate/nmd

Spontaneous muscular dystrophy caused by a retrotransposal insertion in the mouse laminin α2 chain gene

Sylvie Besse[a], Valérie Allamand[a,*], Jean-Thomas Vilquin[a], Zhenlin Li[b], Christophe Poirier[c], Nicolas Vignier[a], Hisae Hori[d], Jean-Louis Guénet[b], Pascale Guicheney[a]

[a]*INSERM U523, Institut de Myologie, IFR 'Coeur Muscle et Vaisseaux' no. 14, Groupe Hospitalier Pitié-Salpêtrière, 47 Boulevard de l'hôpital, 75651 Paris Cedex 13, France*
[b]*Université Denis Diderot, Biologie Moléculaire de la Différenciation, 2 place Jussieu, 75005 Paris, France*
[c]*Unité de Génétique des Mammifères, Institut Pasteur, 25 rue du docteur Roux, 75724 Paris Cedex 15, France*
[d]*Division of Adult Diseases, Medical Research Institute, 2-3-10, Kanda Surugadai, Chiyoda-Ku, Tokyo 101-0062, Japan*

Received 13 June 2002; received in revised form 1 November 2002; accepted 15 November 2002

Abstract

We identified a novel spontaneous mouse model of human congenital muscular dystrophy with laminin α2 chain deficiency, named dy^{Pas}/dy^{Pas}. Homozygous animals rapidly developed a progressive muscular dystrophy leading to premature death. Immunohistological and biochemical analyses demonstrated the absence of laminin α2 chain expression in skeletal muscle. Analysis of the laminin α2 chain cDNA showed the insertion of the long terminal repeat of an intracisternal A-particle gene. In addition, a 6.1 kb insertion composed of retrotransposon elements was identified in the *Lama2* sequence. The dy^{Pas}/dy^{Pas} mouse is thus the first spontaneous mutant with a complete laminin α2 chain deficiency in which the mutation has been identified.

Keywords: Muscular dystrophy; Animal model; Skeletal muscle; Laminin α2 chain; Long terminal repeat; Retrotransposon

1. Introduction

Congenital muscular dystrophy (CMD) is a heterogeneous, autosomal recessive, neuromuscular disorder mainly characterized by hypotonia and contractures at birth, early muscle fiber degeneration, muscle fibrosis and white matter abnormalities [1,2]. Mutations in the *LAMA2* gene encoding the laminin α2 chain lead to a partial or complete deficiency in the expression of the protein in about half of the patients [3]. Mutations resulting in partial deficiency are usually associated with a less severe phenotype [4,5]. The α2 chain is a component of laminin-2 (formerly merosin, [6]) and laminin-4, extracellular matrix proteins formed by the association of three subunits: α, β and γ. The laminin-2 heterotrimer is composed of α2, β1 and γ1 chains and is mainly expressed in the basal lamina of skeletal and cardiac muscle cells, trophoblasts, Schwann cells and cells of mesenchymal origin.

Two spontaneous mutant strains representing animal models for CMD with laminin α2 chain deficiency have been described, namely the *dystrophia muscularis dy/dy* [7] and dy^{2J}/dy^{2J} [8]. The identification of laminin α2 chain deficiency in the *dy* and dy^{2J} mice [9–11] has validated these animals as useful tools for investigating etiopathogenic hypotheses or therapeutic strategies. As in human CMD, the disease is transmitted as an autosomal recessive trait. The murine *Lama2* gene is involved in the pathology of the *dy* and dy^{2J} animals, but to date the exact mutation has been characterized only in the dy^{2J}/dy^{2J} mice: a point mutation in a splice site leads to the formation of a truncated but partially functional α2 chain leading to a more benign phenotype [12,13].

Two knock-out mouse models of laminin α2 chain deficiency have recently been described. The dy^{3k}/dy^{3k} [14] and the dy^{W}/dy^{W} [15] mice develop early and severe clinical signs of muscular dystrophy.

In the present study, we describe a new allele, dy^{Pas}, which spontaneously arose at the Pasteur Institute in Paris. The affected animals were first identified by the presentation of an early and severe phenotype of muscular dystrophy. We thus performed analyses of the mutant at the clinical, histological, biochemical and genetic levels and demonstrated that this strain represented an additional model for human CMD with laminin α2 chain deficiency. Unfortunately, this

* Corresponding author. Tel.: +33-1-4216-5720; fax: +33-1-4216-5700.
E-mail address: v.allamand@myologie.chups.jussieu.fr (V. Allamand).

0960-8966/02/$ - see front matter © 2002 Elsevier Science B.V. All rights reserved.
doi:10.1016/S0960-8966(02)00278-X

description has a historical interest, since the strain is now extinct.

2. Materials and methods

2.1. Mouse strains

The new spontaneous mutation was first observed in 1994 at the animal facility of the Pasteur Institute in Paris (France) among the offspring of a F2 progeny during a linkage experiment. For the allelism assay, heterozygote *+/dy* breeders were obtained from The Jackson Laboratory (Bar Harbor, ME) and crossed with breeders of the new mutant mouse, and the clinical phenotype of the offspring was observed.

2.2. Histological observation and immunofluorescence

The mice were euthanized and sural triceps samples were taken and snap frozen in nitrogen-cooled isopentane without prior fixation. Ten μm thick transverse cryosections were stained with hematoxylin-eosin.

The rabbit polyclonal Ab804 directed against the C-terminal 80-kD domain of the α2 chain was used for immunohistofluorescence experiments [16]. Seven-μm thick transverse cryosections were fixed in ice cold acetone for 10 min. Non-specific binding was blocked using 5% fetal bovine serum (FBS) in phosphate buffered saline (PBS) for 30 min. The sections were then incubated for 2 h at 37°C with Ab804 (1/200 dilution). Following several washes with 0.5% FBS in PBS, sections were incubated for 1 h at 37°C with CY™-conjugated anti-rabbit IgG (1/200 dilution, Jackson Immunoresearch Laboratories, Inc, Baltimore, USA). Muscle sections were then photographed using a Zeiss fluorescence microscope. The exposure time was established using a preparation of wild type mouse muscle.

2.3. Protein detection by immunoblotting

Muscle biopsies were processed as previously described [16]. Briefly, biopsies were homogenized in 50 mM Tris, pH 7.4, and 150 mM NaCl (A buffer), containing protease inhibitors (2 mM PMSF, 4 mg/ml Aprotinin, and 4 mg/ml Leupeptin) using a polytron. Following centrifugation at 14 000 × *g* for 15 min at 4°C, proteins were extracted in the A buffer containing 20 mM EDTA. After subsequent centrifugation, supernatants were solubilized by boiling in an sodium dodecyl sulfate (SDS)-containing loading buffer. Proteins were separated on SDS-PAGE gels, followed by overnight transfer onto nitrocellulose. Immunoblots were incubated with mouse 2D9 antibody directed against the 300 kD fragment of the α2 chain [17] (1/1000 dilution) and detected by chemiluminescence (Pierce ECL, Interchim).

2.4. Genetic mapping

Female $+/dy^{Pas}$ heterozygous carriers were mated with MAI males. F1 mice were test-crossed with heterozygous carriers in order to select F1 MAI/*Pas* inbred heterozygous carriers which were then intercrossed. Liver genomic DNA from 30 affected F2 mice was extracted [18]. The genotype of each animal was determined by non-radioactive polymerase chain reaction (PCR) for a series of polymorphic microsatellites spanning the entire mouse genome [19] and analyzed with the MAPMANAGER software [20].

2.5. Reverse transcription (RT)-PCR and cDNA sequencing

Total RNA was isolated from control and dy^{Pas}/dy^{Pas} skeletal muscle using RNA Plus™ Biophenol (Bioprobe Systems, France). RNA was reverse transcribed using the cDNA synthesis kit (Amersham Pharmacia Biotech) with random hexadeoxynucleotides primers. Eight overlapping primer pairs previously described [10] were used to amplify the full-length α2 chain cDNA [13]. PCR products were subcloned in the PGEM-T vector system (Promega Corporation, Madison, USA) and sequenced on an automated sequencer (ABPrism 377, Applied Biosystems) with the Taq cycle sequencing kit (Perkin Elmer, Applied Biosystems, Foster City) using fluorescent dideoxynucleotides and one of the PCR primers.

2.6. Analysis of genomic DNA

Genomic DNA was isolated from skeletal muscle and tails. For the wild type mouse, primers used for the amplification of DNA were designed from the sequence of the control mouse α2 cDNA covering nucleotides 5002–5276 which correspond to exons 34 and 35 in the human α2 chain cDNA (exXF 5′-GCTGAGAGGACCAACTCCAGAGC-3′; exYR 5′-GCTCCTCCAAGTTTCTCTCTGCTG-3′). These primers were used under the following conditions: 94°C for 1 min, followed by 30 cycles of 94°C for 30 s, annealing and elongation at 68°C for 5 min, and a final step of 68°C for 5 min. PCR products were subcloned and sequenced. For the dy^{Pas}/dy^{Pas} mouse, the primers were designed from the cDNA sequence of the dy^{Pas}/dy^{Pas} mouse (exF: 5′-ACTCCAGAGCAGAATCCTTG-3′; insR: 5′-ATAAGACAGCAGAGAGAAACT-3′). This pair of primers was used under the following conditions: 35 cycles at 94°C for 30 s, annealing at 58°C for 1 min and 72°C for 2 min. DNA was amplified in 50-μl reactions with 80 ng of template DNA using Boerhinger (Boehringer Mannhein GmbH, Germany) *Taq* polymerase. Amplified products were sequenced.

Entire intronic insertion was amplified from the dy^{Pas}/dy^{Pas} mouse with the ExpandTM Long template PCR system kit (Boehringer Mannheim GmbH, Germany) using primers IntF (5′-CGCATTAAGTGAGCAACGA-CAGT-3′) and IntR (5′-GCGAACTACGGACAAAATGC-CAG-3′). PCR-amplification conditions were as follows:

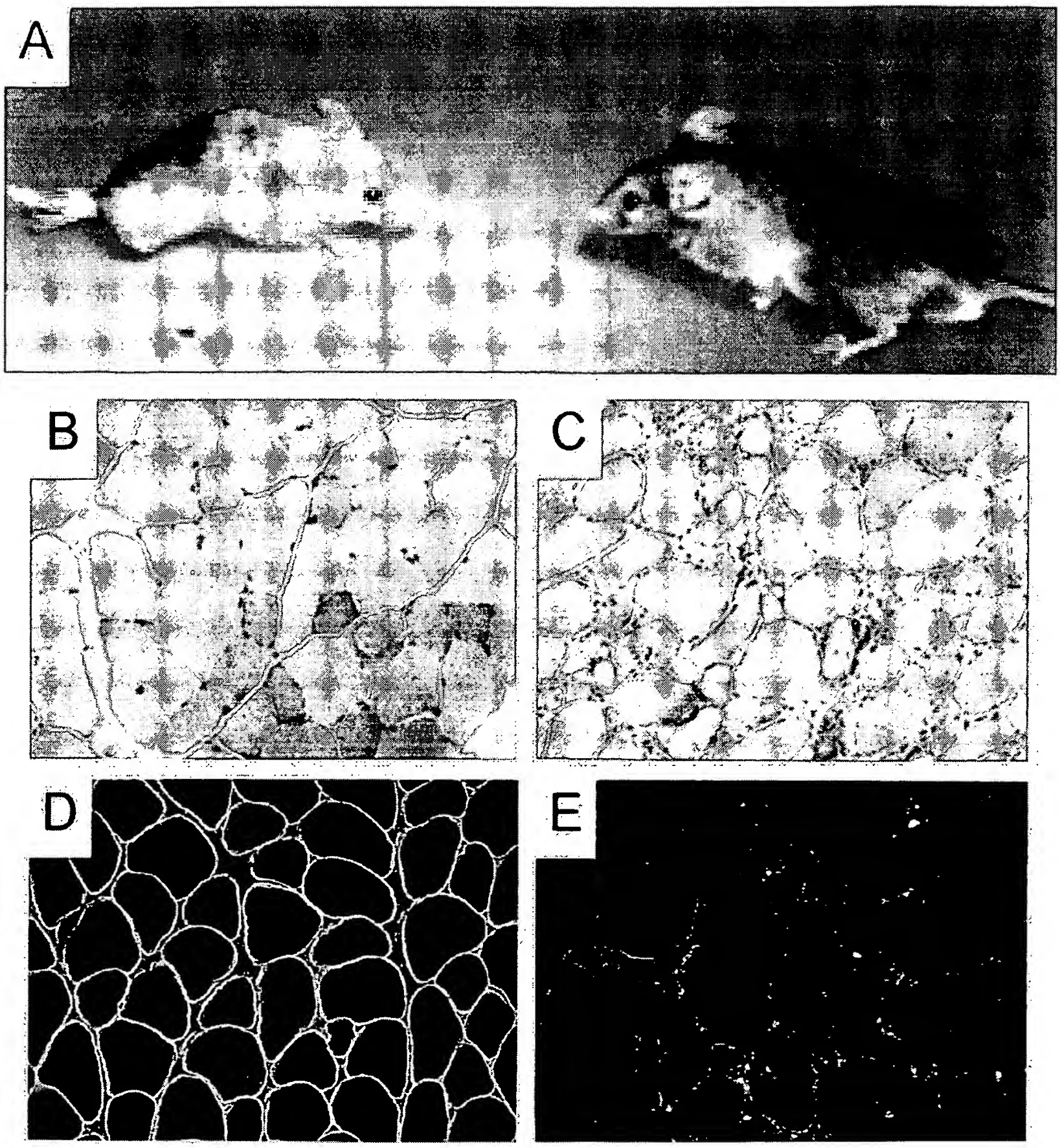


Fig. 1. Phenotypical, histological and immufluorescence characterization of 12 week-old wild-type (A, B, D) and mutant dy^{Pas}/dy^{Pas} (A, C, E) mice. Cryosections of skeletal muscle of normal (B) and homozygous dy^{Pas}/dy^{Pas} (C) mice were stained with hematoxylin and eosin. Muscle fibers are of variable size and some present centrally placed nuclei. The proliferation of connective tissue between the individual muscle fibers is apparent. Original magnification: X160. Localization of the laminin α2 chain was investigated by immunostaining of 7-μm cryosections from normal (D), and dy^{Pas}/dy^{Pas} (E) skeletal muscle with a polyclonal antibody. Normal labeling around muscle fibers was observed only in control mice. In mutant mice, the immunostaining intensity was markedly decreased. Original magnification: × 125.

93°C for 1 min; 93°C for 20 s between each cycle; 30 cycles at 58°C with cycle elongation for more yield of 10 s for the last ten cycles; extension at 68°C for 10 min; and a prolonged elongation time up to 7 min at 68°C. Finally, the PCR products were subcloned and sequenced.

3. Results

3.1. Phenotype of the dy^{Pas}/dy^{Pas} mouse

Animals were derived from a non-inbred agouti back-

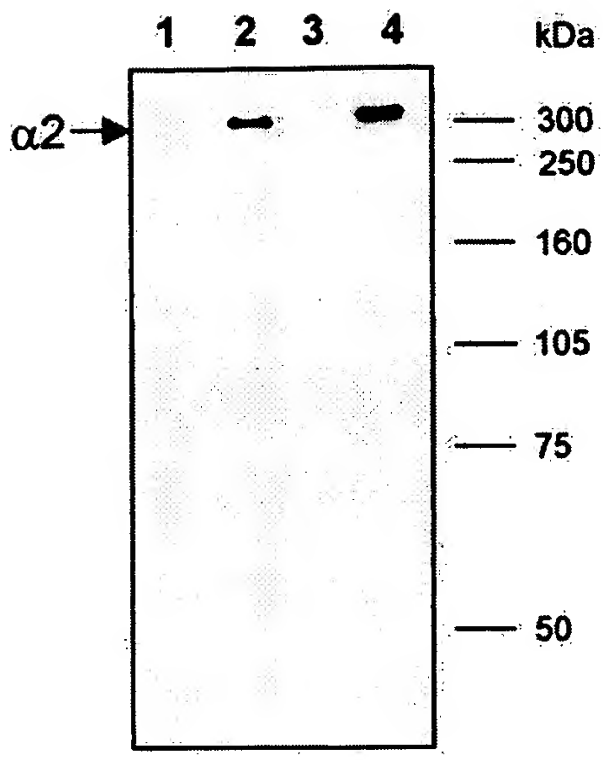


Fig. 2. Immunoblot analysis of control and dy^{Pas}/dy^{Pas} mice using the monoclonal anti-laminin α2 chain. Extracts of skeletal muscle from control mice (2, 4) and dy^{Pas} mice (1, 3). The dy^{Pas}/dy^{Pas} mice lack the 300 kDa segment.

ground. By 2 weeks of age, affected animals were less active than their littermates, and clinical signs became evident within the third week (Fig. 1). When held by the tail, the mice were unable to keep their posterior legs in a tonic flexion position and presented repetitive jerking movements of flexion and extension. They also started waddling and finally developed contractures, joint deformities and weakness. In the subsequent weeks, the posterior legs were immobilized and the mice were dragging themselves when crawling in the cages. The anterior legs were also affected and the mice were supporting themselves on their elbows, and using their anterior legs as oars to crawl in the cages. A kyphoscoliosis developed. The mice remained small, became cachectic and poorly active after weaning, with a body weight never exceeding 12 g. The clinical signs rapidly worsened within the last 3 weeks of life. At this stage the mice were extremely weak, cachectic, they lost mobility control and presented polypnea. The animals never lived longer than 13 weeks. We could never obtain spontaneous breeding of the dy^{Pas}/dy^{Pas} mice.

3.2. Localization of the disease locus

We found linkage of the dy^{Pas} locus with mouse chromosome 10 microsatellite markers placing the mutation between *D10Mit15* and *D10Mit16* [18] in the genomic region harboring the *Lama2* gene. A complementation test with the original allele *dy* was performed by crossing a heterozygous carrier of this mutation and a heterozygous carrier of the $Lama2^{dy}$ mutation. Two out of ten mice from this cross displayed the clinic features of the homozygous affected mice demonstrating the inability of the dy^{Pas} mutation to complement the $Lama2^{dy}$. This test confirmed allelism of the two mutations.

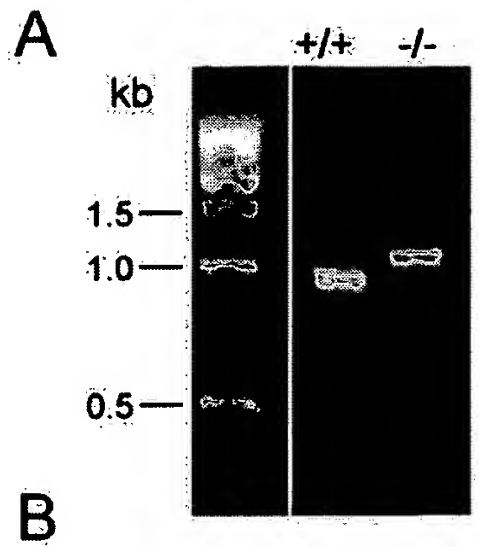
3.3. Histological and immunochemical analysis

The main histopathological changes observed in the homozygous dy^{Pas} mice (Fig. 1C) were variation in fiber size, centrally located nuclei, muscle fiber necrosis, and peri- and endomysial connective tissue proliferation. These lesions are characteristic of congenital muscular dystrophy in mouse and human.

Immunofluorescence localization of the laminin α2 chain in skeletal muscle from wild-type and dy^{Pas}/dy^{Pas} mice was examined using a polyclonal rabbit antibody against the C-terminal portion of the α2 chain. Normal expression of this protein at the basal lamina surrounding each muscle fiber was observed in the control (Fig. 1D) while the immunostaining was absent in the mutant mice (Fig. 1E). A deficiency of the laminin α2 chain in dy^{Pas}/dy^{Pas} skeletal muscle was confirmed by Western blot (Fig. 2). Using the 2D9 antibody, the 300 kDa fragment of α2 chain was detected in normal, but not dy^{Pas}/dy^{Pas} muscle.

3.4. Identification of the dy^{Pas}/dy^{Pas} mutation

Analysis of muscle RT-PCR products covering the laminin α2 chain coding region revealed a larger fragment size in the dy^{Pas}/dy^{Pas} than in the control (1377 versus 1087 bp) (Fig. 3A). Sequencing of the abnormal product revealed a 290 bp insertion at nucleotide position 5113. The insertion showed high homology with a 5′ long terminal repeat (LTR) intracisternal-A particle element (IAP; 99% similarity; GenBank accession no. M59201/M64837/K01235). To further characterize the nature of the rearrangement, we amplified, cloned and sequenced the region from control and mutant DNA with various combination of primers. Alignment of mouse *Lama2* and human *LAMA2* [21] sequences indicated that the 290 bp insertion occurred between exons 34 and 35. Using long range PCR with intronic primers, a 6215 bp insertion was found at position +325 of the intronic sequence between exons 34 and 35 in the dy^{Pas}/dy^{Pas} mouse. This insertion was composed of $(T)_{56}$, a 290 bp sequence corresponding to a 5′ LTR-IAP element and a F-type mouse L1 retrotransposon. The complete sequence is available at GenBank, accession number AJ277888. The two sequences, CAAAG preceded by a pyrimidine-rich region (56 T) and GTGAGT, which flank the inserted IAP-LTR element, fit with the characteristic 3′ and 5′ splicing consensus sequences, YNYAG and GTRAGY, respectively (Fig. 3B). These cryptic sites in the retrotransposon are thought to be responsible for the 290-bp insertion in the transcript. The translational reading frame was thereby shifted from amino acid position 1706 and a stop codon appeared 12 codons later. The abnormal transcript was the main spliced product in skeletal muscle and would induce a putative protein lacking most of domains I and II, and the C-terminal G globular domain.



B

exF

AGAGCAAACAGGACAAGATGCTGAGAGGACCA*ACTCCAGAGCAGAATCCTTG*GAA
GAATTCATTAAAGGGCTTGTCCAGGATGCTGAAGgtatcaaaaaaaaatcatatt
ttaacttttttctcttacataaaaattaaaagaagggattaaagggatttgaata
ggccctataaggaggtgtcatttctaaattcaagttgtaagtgtagaatgcattt
tgtttgacatattcagggtcatgattagcagaaatctatagcctatcgcattaag
tgagcaacgacagttgatgaaataatgattatgggtgtaaaataggaatgaattc
taggacgtttaaataagtcattgaaaacatatggctttgctatttttctgaccag
agctattgaaaattgattcatcttttct**ctttttttttttttttttttttttt**
ttttttttttttttttttttttttttttctgcggcaaagctttatttcttaca
tcttcaggagcaagcgtgtaagaagcaagcgagcaagaagcaagagagagagaaa
acgaaaccccgtccctcttaaggagcattctccttcgcctcggacgtgtcactcc
ctgattggctgcagcccatcggccgagttgacgtcacggggaaggcagagcacaa
gtagtcataagataccttggcacatgcgcagattatttgtttaccacttagaac
acagctgtcagcgccatct*tgtaacggcgaatgtgggc*gcggctcccaacaattg

insR

tgtgagttcctttgtgattgtgttacctcactcaggatgatgctctcca... 7kb...
ttttaaaagttagttttaccattctatttggtacacaaacactgttctttcattt
cttaaaatggttctttaagtacttggtatgagttgcaagag*aacatctggtgctt*

intR

*ggatttggc*tacccgagttcaccagaggct...3..9kb...ttgacaaaatttccctta

tctcctcccatagCCATAAATGAAAAAGCTGTAAAACTAA

Fig. 3. (A) Analysis of control (+/+) and dy^{Pas}/dy^{Pas} (−/−) *Lama2* cDNA by RT-PCR. The mutant fragment was larger than in the control. (B) Part of the nucleotide sequence of dy^{Pas}/dy^{Pas} mouse intron corresponding to the genomic insertion. Part of the mutant sequence is given with the exons in captions, normal intronic in small letters, mutant sequence in small bold letters. The consensus splice sites are boxed, and the abnormally spliced sequence underlined. The sequences corresponding to primers are underlined with dotted lines. The mutant intronic sequence has been deposited in GenBank (accession number: AJ277888).

4. Discussion

The spontaneous mutant mouse described here is the first animal model of human CMD with complete laminin α2 chain deficiency in which the mutation has been identified. In fact, no mutation has been reported in the *dy/dy* dystrophic mouse, which expresses a very low level of laminin-2, and is severely affected. Only a single base mutation that causes an aberrant splicing of the *Lama2* gene has been identified in the allelic dy^{2J}/dy^{2J} mouse, a spontaneous mutant with partial laminin-2 deficiency, with a less severe disease.

Unfortunately, multiple attempts to transfer embryos and expand the colony according to pathogen-free conditions failed, and the strain consequently became extinct. The present description is therefore mainly of historical interest.

The phenotype of the dy^{Pas}/dy^{Pas} mice resembles the *dy/dy* more closely than the dy^{2J}/dy^{2J} mice, with a rapid worsening of clinical signs and a shorter life expectancy. It is worth noting that the dy^{Pas}/dy^{Pas} animals were not of the same genetic background as the *dy/dy* and dy^{2J}/dy^{2J} animals. Although it was reported that varying the genetic background of *dy/dy* animals did not change the clinical status

dramatically [22,23], this difference may nevertheless account for subtle variations between clinical signs, such as susceptibility to ptosis.

Analysis of skeletal muscle histology demonstrated the absence of the laminin α2 chain. The muscle tissue shows a variability of muscle fiber diameter and a connective tissue proliferation.

We have demonstrated that the disruption of the laminin α2 chain gene is caused by a 6.1 kb insertion of LTR-IAP and F-type L1 retrotransposon elements in an intron of the *Lama2* gene. These elements are classically present in multiple copies in mammalian genomes, in which they retrotranspose through an RNA intermediate step. In man, three inherited disorders resulting from exonic L1 retrotransposition have been described: (1) L1 insertion into the factor VIII gene in two patients with haemophilia A [24]; (2) L1 insertion in the dystrophin gene in one patient with Duchenne muscular dystrophy [25]; and (3) L1 insertion in the fukutin gene of one patient with a severe form of Fukuyama-type congenital muscular dystrophy [26]. In mouse, the *spa* model is due to an L1 insertion into the glycine receptor β-subunit gene [27,28]. Although L1 retrotransposons have been involved both in human and murine diseases, IAP elements are present in murine genome in about 2000 copies, but have not yet been identified in the human genome [29–34]. IAP are flanked by 5′- and 3′-LTRs which contain regulatory sequences required for transcription of viral genes. In addition to directing their own transcription, LTRs are able to promote the transcription of adjacent cellular sequences. Mouse mutations have been shown to be caused by IAP element insertions into exonic and intronic sequences that would introduce cryptic splice sites and disrupt normal transcript processing. Several examples have been described, such as: integration into the intron of an actively transcribed immunoglobulin gene [35]; insertion in the tyrosinase promoter [36]; insertion into *Eya1* gene [37]. Interestingly, in the dy^{Pas}/dy^{Pas} mouse, the insertion is composed of both L1 and LTR-IAP elements at the genomic level, but only the 5′-LTR sequence of the IAP is present at the mRNA level. By comparison between the cDNA and genomic sequences, we concluded that the insertion in the dy^{Pas}/dy^{Pas} mouse creates cryptic sites and allows the splicing of LTR element.

The 290-bp insertion produces an aberrant transcript, which would result in a putative truncated protein devoid of domain I + II and the C-terminal G domain. In laminin-2, domain I + II is involved in the formation of the coiled-coil long arm. In muscle, the laminin α2 chain attaches through the globular G domain with α-dystroglycan and integrin α7β1 and provides a link between the extracellular matrix and sarcolemmal cytoskeleton [38]. Histopathological changes have been identified in both *dy/dy* and dy^{2J}/dy^{2J} basement membranes [10,12]. From these observations, one may hypothesize that the absence of the laminin α2 chain in dy^{Pas}/dy^{Pas} mice leads to disruption of the basal lamina and to weakened anchorage of muscle fibers to extracellular matrix, leading to muscular dysfunction.

Acknowledgements

The authors are grateful to Dr E. Engvall for generously providing the AB804 antibody, and to Dr F. Tomé for his comments concerning the manuscript. This work was supported by the Institut National de la Santé et de la Recherche Médicale (INSERM), the Association Française contre les Myopathies (AFM, France), the Fondation pour la Recherche Médicale (to V.A.) and the European Community (Contract No QLG1-1999-00870).

References

[1] Banker BQ. The congenital muscular dystrophies. In: Engel AG, Franzini-Amstrong C, editors. Myology, 2, 2nd ed.. New York: McGraw-Hill, 1994. pp. 1275–1289.
[2] Fardeau M. Congenital myopathies. In: Mastaglia FL, Detchant LWO, editors. Skeletal muscle pathology, 2nd ed.. Edinburgh: Churchill Livingstone, 1992. pp. 237–281.
[3] Helbling-Leclerc A, Zhang X, Topaloglu H, et al. Mutations in the laminin α2-chain gene (LAMA2) cause merosin-deficient congenital muscular dystrophy. Nat Genet 1995;11:216–218.
[4] Nissinen M, Helbling-Leclerc A, Zhang X, et al. Substitution of a conserved cysteine-996 in a cysteine-rich motif of the laminin α2-chain in congenital muscular dystrophy with partial deficiency of the protein. Am J Hum Genet 1996;58:1177–1184.
[5] Allamand V, Sunada Y, Salih MAM, et al. Mild congenital muscular dystrophy in two patients with an internally deleted laminin α2-chain. Hum Mol Genet 1997;6:747–752.
[6] Burgeson RE, Chiquet M, Deutzmann R, et al. A new nomenclature for the laminins. Matrix Biol 1994;14:209–211.
[7] Michelson AM, Russell E, Harman PG. Dystrophia muscularis: a hereditary primary myopathy in the mouse. Proc Natl Acad Sci USA 1955;41:1079–1084.
[8] Meier H, Southard JL. Muscular dystrophy in the mouse caused by an allele at the *dy* locus. Life Sci 1970;9:137–144.
[9] Arahata K, Hayashi YK, Koga R, et al. Laminin in animal models for muscular dystrophy: defect of laminin M in skeletal and cardiac muscles and peripheral nerve of the homozygous dystrophic *dy/dy* mice. Proc Jpn Acad 1993;69:259–264.
[10] Sunada Y, Bernier SM, Kozak CA, Yamada Y, Campbell KP. Deficiency of merosin in dystrophic *dy* mice and genetic linkage of the laminin M chain gene to the *dy* locus. J Biol Chem 1994;269:13729–13732.
[11] Xu H, Christmas P, Wu X-R, Wewer UM, Engvall E. Defective muscle basement membrane and lack of M-laminin in the dystrophic *dy/dy* mouse. Proc Natl Acad Sci USA 1994;91:5572–5576.
[12] Xu H, Wu XR, Wewer UM, Engvall E. Murine muscular dystrophy caused by a mutation in the laminin α2 (LAMA2) gene. Nat Genet 1994;8:297–302.
[13] Sunada Y, Bernier SM, Utani A, Yamada Y, Campbell KP. Identification of a novel mutant transcript of laminin α2 chain gene responsible for muscular dystrophy and dysmyelination in dy^{2J} mice. Hum Mol Genet 1995;4:1055–1061.
[14] Miyagoe Y, Hanaoka K, Nonaka I, et al. Laminin a2 chain-null mutant mice by targeted disruption of the LAMA2 gene: a new model of merosin (laminin 2)-deficient congenital muscular dystrophy. FEBS Lett 1997;415:33–39.
[15] Kuang W, Vachon PH, Liu L, Loechel F, Wewer UM, Engvall E.

Merosin-deficient congenital muscular dystrophy. Partial genetic correction in two mouse models. J Clin Invest 1998;102:844–852.
[16] Ehrig K, Leivo I, Argraves WS, Ruoslahti E, Engvall E. Merosin, a tissue-specific basement membrane protein, is a laminin-like protein. Proc Natl Acad Sci USA 1990;87:3264–3268.
[17] Hori H, Kaa T, Mizuta T, Yamaguchu N, Liu Y, Nagai Y. Human laminin M chain: epitope analysis of its monoclonal antibodies by immunoscreening of cDNA clones and tissue expression. J Biochem 1994;116:1212–1219.
[18] Blin N, Stafford DW. A general method for isolation of high molecular weight DNA from eukaryotes. Nucleic Acids Res 1976;3:2303–2308.
[19] Dietrich WF, Miller J, Steen R, et al. A comprehensive genetic map of the mouse genome. Nature 1996;380:149–152.
[20] Manly KF. A Macintosh program for storage and analysis of experimental genetic mapping data. Mamm Gen 1993;4:303–313.
[21] Zhang X, Vuolteenaho R, Tryggvason K. Structure of the human laminin a2-chain gene (*LAMA2*), which is affected in congenital muscular dystrophy. J Biol Chem 1996;271:27664–27669.
[22] Totsuka T, Watanabe K. Some evidence for concurrent involvement of the fore- and hindleg muscles in murine muscular dystrophy. Jikken Dobutsu 1981;30:465–470.
[23] Meier H, MacPike AD. Myopathies caused by three mutations of the mouse. J Hered 1977;68:383–385.
[24] Kazazian HHJ, Wong C, Youssoufian H, Scott AF, Phillips DG, Antonarakis SE. Haemophilia A resulting from de novo insertion of L1 sequences represents a novel mechanism for mutation in man. Nature 1988;332:164–166.
[25] Narita N, Nishio H, Kitoh Y, et al. Insertion of a 5′ truncated L1 element into the 3′ end of exon 44 of the dystrophin gene resulted in skipping of the exon during splicing in a case of Duchenne muscular dystrophy. J Clin Invest 1993;91:1862–1867.
[26] Konda-Iida E, Saito K, Tanaka H, et al. Molecular genetic evidence of clinical heterogeneity in Fukuyama-type congenital muscular dystrophy. Hum Genet 1997;99:427–432.
[27] Kingsmore SF, Giros B, Suh D, Bieniarz M, Caron MG, Seldin MF. Glycine receptor beta-subunit gene mutation in spastic mouse associated with LINE-1 element insertion. Nat Genet 1994;7:143–148.
[28] Mulhardt C, Fischer M, Gass P, et al. The spastic mouse: aberrant splicing of glycine receptor beta subunit mRNA caused by intronic insertion of L1 element. Neuron 1994;13:1003–1015.
[29] Kuff EL, Feenstra A, Lueders K, et al. Intracisternal A-particle genes as movable elements in the mouse genome. Proc Natl Acad Sci USA 1983;80:1992–1996.
[30] Kuff EL, Lueders KK. The intracisternal A-particle gene family: structure and functional aspects. Adv Cancer Rev 1988;93:183–276.
[31] Cole MD, Ono M, Huang RC. Terminally redundant sequences in cellular intracisternal A-particle genes. J Virol 1981;38:680–687.
[32] Lueders K, Kuff EL. Sequences associated with intracisternal A particles are reiterated in the mouse genome. Cell 1977;12:963–972.
[33] Lueders K, Kuff EL. Intracisternal A-particle genes: identification in the genome of Mus musculus and comparison of multiple isolates from a mouse gene library. Proc Natl Acad Sci USA 1980;77:3571–3575.
[34] Ono M, Cole MD, White AT, Huang RC. Sequence organization of cloned intracisternal A particle genes. Cell 1980;21:465–473.
[35] Hawley RG, Shulman MJ, Murialdo H, Gibson DM, Hozumi N. Mutant immunoglobulin genes have repetitive DNA elements inserted into their intervening sequences. Proc Natl Acad Sci USA 1982;79:7425–7429.
[36] Wu M, Rinchik EM, Wilkinson E, Johnson DK. Inherited somatic mosaicism caused by an intracisternal A particle insertion in the mouse tyrosinase gene. Proc Natl Acad Sci USA 1997;94:890–894.
[37] Johnson KR, Cook SA, Erway LC, et al. Inner ear and kidney anomalies caused by IAP insertion in an intron of the Eya1 gene in a mouse model of BOR syndrome. Hum Mol Genet 1999;8:645–653.
[38] Colognato H, Yurchenco PD. The laminin alpha2 expressed by dystrophic dy(2J) mice is defective in its ability to form polymers. Curr Biol 1999;9:1327–1330.